CYBERLAW

BY CIJIC

DIREITO: A PENSAR TECNOLOGICAMENTE

EDIÇÃO N.º XIV

Setembro de 2026

CIJIC
Centro de Investigação Jurídica do Ciberespaço
da Faculdade de Direito de Lisboa

FACULDADE DE DIREITO
UNIVERSIDADE DE LISBOA

Iuris
Instituto de
Investigação
Interdisciplinar



CYBERLAW

BY CIJIC

EDIÇÃO N.º XIV - SETEMBRO DE 2026

REVISTA CIENTÍFICA SOBRE CYBERLAW DO CENTRO DE INVESTIGAÇÃO JURÍDICA DO CIBERESPAÇO - CIJIC

FACULDADE DE DIREITO DA UNIVERSIDADE DE LISBOA

X ANIVERSÁRIO



CYBERLAW

BY CIJIC

DIRETOR REVISTA: NUNO TEIXEIRA CASTRO

EDITORIAL: AFONSO DE FREITAS DANTAS – MARGARIDA FERREIRA - EUGÉNIO ALVES DA SILVA

COMISSÃO CIENTÍFICA DA REVISTA:

- ALFONSO GALAN MUÑOZ
- ANGELO VIGLIANISI FERRARO
- ANTÓNIO R. MOREIRA
- DANIEL FREIRE E ALMEIDA
- ELLEN WESSELINGH
- FRANCISCO MUÑOZ CONDE
- MANUEL DAVID MASSENO
- MARCO ANTÓNIO MARQUES DA SILVA
- MARCOS WACHOWICZ
- ÓSCAR R. PUCCINELLI
- RAQUEL A. BRÍZIDA CASTRO

CIJIC: CENTRO DE INVESTIGAÇÃO JURÍDICA DO CIBERESPAÇO

ISSN 2183-729

CYBERLAW

BY CIJIC

ÍNDICE:

I - **INTELIGÊNCIA ARTIFICIAL E ATIVIDADE JURISDICIONAL: ENTRE A EPISTEMOLOGIA JURÍDICA E OS DESAFIOS DE GOVERNAÇÃO.**

NUNO COELHO ---------- págs. 25 - 82

II- **INTELIGÊNCIA ARTIFICIAL (IA) E JUÍZES - PISTAS PARA O FUTURO.**

JOSÉ JOAQUIM FERNANDES OLIVEIRA MARTINS -- ---------- págs. 83 - 111

III- ***IUS EX MACHINA?*** **ENTRE UMA ADMINISTRAÇÃO PÚBLICA "ARTIFICIALMENTE INTELIGENTE" E A TUTELA DOS DIREITOS FUNDAMENTAIS DOS ADMINISTRADOS.**

LUÍS MANUEL PICA ---------- págs. 112- 135

IV - **O RISCO OCULTO DA INTELIGÊNCIA ARTIFICIAL: COMO A AUTOMATIZAÇÃO PODE DEGRADAR A CAPACIDADE COGNITIVA, LABORAL, ORGANIZACIONAL E SECURITÁRIA.**

FERNANDO CAVALEIRO ÂNGELO ---------- págs. 136 - 158

CYBERLAW

BY CIJIC

V - **O PROFISSIONAL DE SAÚDE NA ERA DIGITAL: RESPONSABILIDADE, AUTONOMIA E CARGA ADMINISTRATIVA.**

DANIELA MARQUES DIAS & CARLA BARBOSA ---------- págs. 159 - 187

VI - **O ENSINO DO DIREITO NA ERA DA INTELIGÊNCIA ARTIFICIAL: A ÉTICA E A DECISÃO HUMANA COMO FUNDAMENTOS DO USO PEDAGÓGICO DA IA, A PARTIR DA EXPERIÊNCIA DA HISTÓRIA DO DIREITO.**

RAIMUNDO CHAVES NETO ---------- págs. 188 - 220

VII - **A REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL EM PORTUGAL: CONTRIBUTO DE LEGE FERENDA PARA UMA LEI DE EXECUÇÃO DO REGULAMENTO (UE) 2024/1689 (EU AI ACT).**

FERNANDA DUARTE MARQUES ---------- págs. 221 - 270

VIII - **DOS BENS INFORMÁTICOS FUNDAMENTAIS À HARMONIZAÇÃO DAS POLÍTICAS DE PROMOÇÃO DA CIBERSEGURANÇA E PREVENÇÃO DO CIBERCRIME.**

PEDRO DIAS VENÂNCIO ---------- págs. 271 - 291

CYBERLAW

BY CIJIC

IX - **O NOVO ART. 8.º-A DA LEI DO CIBERCRIME: CONSIDERAÇÕES BREVES SOBRE O HACKING ÉTICO EM PORTUGAL, DESDE UMA PERSPETIVA SIMULTANEAMENTE TÉCNICA E JURÍDICA.**

PEDRO MARTINS PROENÇA & MANUEL DAVID MASSENO -------- págs. 292 - 317

X - ***O CYBERCRIME-AS-A-SERVICE (CAAS).***

DUARTE RODRIGUES NUNES ---------- págs. 318 - 346

XI - **PARA ALÉM DO AVATAR DIGNIDADE HUMANA, IDENTIDADE E COMUNIDADE DIGITAL NA VIDA EXTRAORDINÁRIA DE IBELIN.**

NUNO POIARES & ANA RITA PULGA ---------- págs. 347 - 375

XII - **OS CÓDIGOS DE CONDUTA EM MATÉRIA DE PROTEÇÃO DE DADOS.**

JOEL A. ALVES ---------- págs. 376 - 392

XIII - **FROM COMPLIANCE TO INTEGRABILITY - DATA PROTECTION IN M&A TRANSACTIONS.**

DIOGO GUERREIRO DUARTE ---------- págs. 393 - 441



NOTAS SOBRE A XIV EDIÇÃO:

Nenhuma tecnologia, como a Inteligência Artificial, é neutra quando interfere decisivamente com a vida das pessoas. Dissemo-lo já, no limiar desta edição. Mas uma convicção (pré)anunciada não é uma convicção demonstrada. Na décima quarta edição da «***Cyberlaw by CIJIC***», dez anos após termos iniciado este nosso percurso de publicações, procuraremos trazer uma demonstração, artigo a artigo, daquilo que a presente edição vos (nos) quer provocar.

Aqui, desde o início, a pessoa humana é o fim de todas as coisas. Mas dizer que a pessoa é o fim de todas as coisas é dizer também que tudo aquilo que sobre ela se constrói - relações, instituições, decisões, tecnologias - só pode cumprir a sua função se preservar o espaço de confiança em que a dignidade humana se realiza. A dignidade humana sem confiança é uma promessa por cumprir. A confiança sem dignidade é apenas um sistema destinado a errar, porque inevitavelmente falhará.

Cremos, por isso, que não há dignidade humana que resista sem confiança. Assim como não haverá confiança que nasça sem dignidade. Uma não vem antes ou depois da outra: nascem

juntas. Ou não nascem. E é sobre este nascimento comum, e sobre tudo aquilo que o pode reforçar ou fragilizar, que procuraremos desvelar «*o tanto*» daquilo que esta edição, especial, temática, tem para nos dizer.

Em dez anos - e parabéns a nós pela efeméride - as histórias, as experiências, as nossas perspectivas, tudo à nossa volta foi mudando. Sobretudo com a presença crescente da IA no nosso quotidiano colectivo. Partindo deste pressuposto, e dando início ao nosso excurso, comecemos, então, pela sala de audiências. Não porque seja ali que a confiança nasce - seria insolente pensá-lo. A confiança é mais antiga do que qualquer tribunal, tão velha quanto a própria vida em comunidade. Comecemos aqui porque é na sala de audiências que a sua ausência (da confiança) comporta um dos maiores riscos para a nossa estabilidade social.

Neste conspecto, merece atenção o estudo da Direção-Geral da Política de Justiça[1] sobre as percepções de satisfação dos utentes da Justiça-Tribunais, referente a 2025. Entre os indicadores avaliados, a imparcialidade do juiz foi o mais bem classificado. Também as dimensões relativas ao juiz responsável pelo processo e à lealdade obtiveram classificações médias superiores. Estes factos merecem atenção: a confiança na justiça não se constrói apenas através da imagem pública das instituições, mas também das percepções que os cidadãos formam sobre o seu funcionamento, na medida em que o escrutínio da experiência real é mais persistente do que a imagem que circula sem ela.

É precisamente esta confiança, a que se ganha (ou perde) caso a caso, no dia-a-dia, que a IA vem testar no imediato. Quando entra no processo de julgar, algo mais visceral mexe do que a mera introdução de uma ferramenta. Pode reconfigurar-se o próprio modo como conhecemos e interpretamos o Direito. Distinguir três níveis de intervenção tecnológica na função de julgar - a gestão do processo, o apoio à decisão e a reflexão crítica sobre ela – ajuda a perceber o que pode ser delegado a operações de classificação e de ordenação e o que continua a exigir um juízo humano, atento às circunstâncias e à singularidade de cada caso concreto. A fronteira decisiva está, precisamente, naquilo que pode ser automatizado sem que se perca a compreensão da Pessoa, o reconhecimento do Outro, e a percepção do conflito submetido a julgamento.

1(Disponível em: https://dgpj.justica.gov.pt/Noticias-da-DGPJ/DGPJ-realiza-estudo-sobre-satisfacao-nos-tribunais-2025)

Desta fronteira depende se continuamos a falar de justiça aumentada ou se começamos já a falar de justiça substituída. O que, indubitavelmente, nos coloca ante o problema de percebermos que a formação dos magistrados deixa de poder ser tratada como uma questão administrativa ou de gestão e passa a ser também uma questão de legitimidade.

Descendo o princípio ao dia-a-dia de quem julga, com este quadro em mente, no plano concreto da actividade jurisdicional, as «*Recomendações no uso de inteligência artificial» do Conselho Superior da Magistratura*»[2] fazem sentido e são assertivas no conteúdo, embora cheguem num momento em que a tecnologia já avançou consideravelmente. Porém, sendo recomendações - ainda que acompanhadas por mecanismos de supervisão - e não uma regra vinculativa de natureza disciplinar, será legítimo perguntar que consequências jurídicas poderá ter a utilização de IA por um juiz em desconformidade com essas orientações. Estará em causa alguma irregularidade disciplinar? Ou poderá essa utilização, sobretudo quando não é transparente nem susceptível de escrutínio, comprometer a legitimidade da decisão perante quem nela confia sem dispor de meios para o auditar? Nem toda a utilização *não autorizada* compromete, *per se*, a validade da decisão humana. Como resolver esta tensão?

A questão já não é uma mera hipótese futura. Há tribunais, em vários países, a testar sistemas de apoio à decisão e magistraturas a escrever, com premência, as suas primeiras directrizes de uso. (O alvoroço costuma ser péssimo conselheiro.) Se a técnica transformar as condições em que o Direito é aplicado, caberá ao Direito assegurar que não transforma, com elas, a própria natureza humana do acto de fazer justiça. Não obstante, esta raia tecnológica reproduz uma outra questão a montante da própria legitimidade do resultado do sistema. Uma questão que incomoda: *No fim do dia, que justiça queremos reconhecer, enquanto comunidade, como Iustitia*?

Todos compreendemos o brocardo segundo o qual "*a justiça é cega*". A imparcialidade veda a acepção de pessoas, mas não dispensa o julgador de ver a pessoa concreta. A venda da *Iustitia* procura exprimir uma imparcialidade do Juiz e a igualdade de todos perante a lei, naquele sentido de recusa de que a decisão dependa de quem seja a pessoa julgada. A venda não retira à justiça o dever de olhar para a pessoa concreta, cuja dignidade (art.º 1.º da Constituição) faz da decisão judicial um acto dirigido a alguém, e não um mero (autómato de) tratamento de casos. Porque julgar não é apenas classificar factos, identificar normas e

2 (Disponível em: https://csm.org.pt/wp-content/uploads/2026/05/Recomendacoes_Uso_IA.pdf)

produzir uma conclusão. É compreender o caso concreto, ponderar as circunstâncias, reconhecer a efectiva dimensão humana do conflito e assumir, perante o Outro, a responsabilidade pela decisão que se toma.

É também por isto que a presença humana não pode ser um mero resíduo, uma formalidade, sobrevivente de um modelo jurídico anterior à tecnologia. É, e terá de continuar a sê-lo, parte definidora da própria ontologia do acto de julgar. A inteligência artificial pode assistir o juiz, pode ampliar a sua capacidade de pesquisa, ordenar informação, identificar padrões ou até desafiar o seu primeiro raciocínio. Mas não poderá tornar-se o «*sujeito invisível*», qual *divindade omnipresente*, de uma decisão que continuará a ser apresentada como uma decisão de justiça. Porque a justiça não é apenas feita para pessoas. É feita de pessoas, para pessoas, pelas pessoas.

Este (mesmo) problema (re)aparece na Administração Pública, deslocado do juiz para o funcionário - com consequências talvez mais profundas por ausência da solenidade e do escrutínio mais típicos de um tribunal. Por exemplo, quando etapas do procedimento são substituídas por regras automáticas, a tramitação transforma-se num sistema híbrido onde a vontade humana se dilui e a autoria do acto se vai fragmentando. Entre quem programou o sistema, quem o configurou, quem o utilizou e quem, no fim, assina uma decisão - que até talvez já não seja inteiramente sua - a responsabilidade torna-se menos evidente. Institutos que julgávamos consolidados - a fundamentação, o direito de ser ouvido antes de uma decisão, a própria identificação do autor de um acto - têm de ser revisitados, não por moda académica, mas porque a decisão automatizada os põe à prova, constantemente, em condições que não haviam sido originariamente ponderadas.

Num mundo hiperconectado, Portugal não escapa a este movimento. A Agenda Nacional de Inteligência Artificial[3], aprovada em Janeiro de 2026 e assente num investimento anunciado superior a quatrocentos milhões de euros até 2030, parte de um estudo que estima em oito por cento as funções do sector público com potencial de automação por inteligência artificial generativa, e em sessenta e seis por cento aquelas com potencial médio. Ainda que tais percentagens não correspondam concretamente a percentagens práticas de funções que poderão ser automatizadas, estes números acabam por conceder à «*procedimentalização algorítmica*» uma escala e uma dimensão estratégica. Quanto mais o cidadão interage com

3 (Disponível em: https://digital.gov.pt/pt/estrategia-digital/agenda-nacional-de-ia)

uma Administração mediada por algoritmos de triagem ou de risco, menos visível se torna quem, de facto, decide. E mais urgente se torna que as garantias do procedimento não cedam ao argumento, sempre sedutor, da *eficiência mecânica*. A Administração Pública tem de estar próxima do cidadão. Essa proximidade é fundamental para a confiança nas instituições. E é também uma das fronteiras que nos permite distinguir um sistema pensado para as pessoas de uma possível *Algoritmocracia*.

Esta, porém, não se proclama forçosamente pela substituição, mais ou menos explícita, da decisão humana. Começa muito antes, quando a conveniência da automatização torna menos perceptível aquilo que vamos delegando, que vamos deixando de fazer por nós próprios.

Desenganem-se os que imaginam que a passagem de uma *IA experimental* para uma IA sistemicamente integrada acarreta *apenas* ganhos de produtividade. Pode também acentuar as desigualdades, amplificar a desinformação e favorecer uma certa lassidão cognitiva, que se vai instalando, sobretudo quando a literacia digital não acompanha a velocidade da transformação tecnológica. No trabalho, na educação, na segurança e na defesa - onde os drones e as operações de manipulação e interferência de informação estrangeiras, conhecidas pela sigla FIMI (*Foreign Information Manipulation and Interference*), já fazem parte do vocabulário estratégico - a pergunta ecoa sob formas diferentes: a tecnologia complementa o pensamento crítico ou cria, pela comodidade que oferece (e como nós humanos gostamos de comodidade!) as condições suficientes para que deixemos de o exercer?

Quid juris, hoje?

O conflito desvela-se (também) na saúde (digital). A automatização em curso não dispensa o profissional de saúde da responsabilidade que lhe compete. Reconfigura-a, mas não a apaga. A tendência para confiar cegamente na sugestão do sistema - *automation bias* - é talvez um dos riscos mais subtis que a era digital introduz na prática clínica, precisamente por não se percepcionar como negligência quando acontece. Como possível solução, a resposta não passará necessariamente por mais tecnologia, mas por mais rigor. Verificação crítica, documentação cuidada e autonomia profissional preservada. É esta cadeia de responsabilidade, e não a confiança (cega ou, pelo menos, turvada) no *divino algoritmo*, que protegerá tanto o doente como o próprio profissional num eventual litígio futuro.

Como já assinaláramos, esta tensão extravasa o plano da *mera hipótese* também no plano nacional. A Comissão para a Inteligência Artificial da Ordem dos Médicos[4] analisou, recentemente, a aplicação do Regulamento Europeu da Inteligência Artificial ao sector nacional da saúde, num contexto em que projectos-piloto em curso no Serviço Nacional de Saúde - na triagem, na avaliação da fragilidade, no uso racional de antibióticos e na imagiologia médica - exigem precisamente supervisão crítica, *transparência, explicabilidade e validação contínua* e não a confiança acrítica que o *automation bias* parece sugerir. Daqui emerge, aliás, um paradoxo, uma crise que a própria promessa da desburocratização revela contra si mesma: uma carga administrativa mal calibrada compromete a relação terapêutica pelo excesso. Porém, desburocratizar sem pensar pode comprometê-la ainda mais pela desresponsabilização que nasce da ausência de registo. Daremos, pois, palco ao «*Projecto SafeCaring*», que propõe um modelo em que a tecnologia apoia sem substituir, apontando para um equilíbrio que nos parece essencial: aquilo que se automatiza deve libertar tempo para o cuidado, para a proximidade humana em momento de evidente fragilidade do doente. Jamais substituir a relação humana de que o cuidado depende.

Mais conhecimento e mais informação. Mais proximidade humana. E uma viagem, enquanto comunidade, de progresso. Sob pena de as instituições em que confiamos, ao longo do tempo, perderem a sua legitimidade. Nada disto, porém, é inteiramente novo. Já em 1990, no último volume da sua trilogia sobre as transformações do poder - «*Powershift*», editado em Portugal como «*Os Novos Poderes*» - Alvin Toffler descrevia o conhecimento como um dos principais amplificadores da força e da riqueza, mostrando como estas próprias formas de poder passavam a depender de quem controlasse a informação e o conhecimento. Bastaria trocar os nomes que Toffler retratou pelos que hoje ocupam manchetes para perceber que o diagnóstico não feneceu: tem-se, ao invés, (re)confirmado. A disputa pelo controlo do conhecimento não é novidade do nosso tempo. É apenas um seu prolongamento mais recente.

É inegável que o conhecimento é poder. E onde há poder, cabe-nos submetê-lo a limites, a responsabilidade e a controlo, através de um Direito que tem a pessoa como fim em si mesma. É por isso que importa olhar para os lugares onde o conhecimento, tradicionalmente, se transmite, se aprende, se questiona e se transforma em capacidade de pensamento crítico, porque é aí que se forma quem há-de exercer, e limitar, esse poder.

---

4 (Disponível em: https://www.actamedicaportuguesa.com/revista/index.php/amp/article/view/24588)

O ensino é um desses lugares - e a entrada da Inteligência Artificial (também) nas faculdades de Direito torna particularmente visível a tensão entre informação e conhecimento. Os modelos de linguagem colocam ao alcance de qualquer estudante, através de uma simples pesquisa, instrumentos capazes de redigir textos jurídicos, sintetizar doutrina, traduzir fontes históricas e simular argumentação. O paradoxo não reside apenas na abundância de recursos, mas na dificuldade de distinguir a resposta produzida do conhecimento efectivamente adquirido. O desafio é, por isso, pedagógico e ético: integrar a tecnologia sem comprometer aquilo que o processo de aprendizagem deve formar, *ie*., a capacidade humana de pensar, fundamentar e discernir.

Um dos contributos desta edição enfrenta esta questão a partir de um terreno particularmente revelador: o Direito Romano e a História do Direito Português. Disciplinas de fontes, de contexto e de juízo, mostram como a tradição histórico-jurídica pode ajudar a compreender os limites dos modelos de linguagem, expondo as suas alucinações e os seus anacronismos, e devolvendo o estudante às fontes que importam interpretar. A experiência pedagógica apresentada não propõe, pois, a resistência à tecnologia, mas a sua integração criteriosa, subordinada à ética e à decisão humana. Recorda-nos, igualmente, que a integridade académica não se defende necessariamente multiplicando a vigilância, agora também automatizada: exige, antes, um desenho pedagógico e avaliativo que torne visíveis o percurso, a compreensão e, acima de tudo, o juízo de quem aprende.

Há, aliás, nesta reflexão, uma advertência que ultrapassa as fronteiras da sala de aula. A evocação da Lei das Citações de 426, na leitura proposta pelo artigo, recorda-nos que a tentativa de disciplinar a decisão através da autoridade contabilizada das fontes antecede em muitos séculos os actuais sistemas algorítmicos. O problema não está em recorrer a instrumentos que auxiliem a decidir, mas em permitir que o mecanismo de auxílio se converta no substituto do juízo. Formar para a era da Inteligência Artificial será, também, preservar a capacidade de reconhecer esta diferença. Porque, paralelamente à discussão sobre os limites que devemos impor às máquinas, importa assegurar que não abdicamos de formar as pessoas chamadas a compreendê-las, a utilizá-las e a responder pelas suas consequências.

Ao poder do conhecimento junta-se o poder tecnológico, e a questão dos seus limites não se esgota na formação de quem o utiliza. Estende-se também à responsabilidade de quem o desenvolve e à capacidade de antecipar, e sobretudo de controlar, as consequências dos

sistemas que cria. É neste plano que o debate sobre a Inteligência Artificial ganha uma dimensão particularmente inquietante: a de investigadores e responsáveis pelo seu desenvolvimento que, conhecendo por dentro as capacidades e as limitações destes sistemas, nos alertaram para riscos que consideram insuficientemente sujeitos a limites, a responsabilidade e a controlo.

Este risco, diríamos sistémico, que acompanha a sociedade actual, marcada por sucessivas crises, ganhou nas últimas semanas rostos que o discurso académico raramente revela: o de quem constrói a tecnologia e, a certa altura, soa um alarme sobre ela própria. Recentemente, um investigador, Jacob Coxon que passara anos em investigação de pré-treino em dois dos principais laboratórios do sector demitiu-se, de um deles, alertando publicamente para aquilo que entendia ser uma resposta insuficiente ao risco de uma inteligência artificial capaz de se *auto-aperfeiçoar*. Evan Hubinger, responsável pela área de alinhamento de uma dessas empresas, juntou-se-lhe, publicamente, estimando, a título pessoal, em mais de dez por cento de probabilidade um cenário de extinção humana na próxima década. Posteriormente, Dario Amodei, por seu turno, tem defendido a necessidade de ponderarmos os riscos associados ao ritmo de desenvolvimento e de reforçar a fiscalização externa.

Ainda assim, nem todos partilham deste nível, mediatizado, *de alarme*. Há quem sustente que uma espécie geneticamente diversa e globalmente dispersa não se deixa extinguir com facilidade, por mais capaz que seja *o sistema* em causa. O debate não se esgota, de resto, na oposição entre previsões catastróficas e confiança no progresso. E que se recuse a fechar não é incidental, é estrutural. Aliás, reforça a dificuldade de anteciparmos as consequências de tecnologias em constante mutação e evolução. Jonathan Zittrain, professor em Harvard, tem vindo a propôr uma tríade[5] para organizar as vozes que hoje disputam o futuro da inteligência artificial, e, que ajude a perceber *porquê*: os idealistas, que nela vêem o caminho para uma espécie de abundância material e temporal; os alarmistas, que temem efeitos irreversíveis, nalguns casos existenciais, sobre as nossas instituições e as nossas vidas; e os cépticos, para quem a preocupação com futuros imaginados, utópicos ou apocalípticos, desvia a atenção dos problemas reais, mais modestos, que a tecnologia já hoje coloca sob pressão. O próprio Zittrain reconhece que as fronteiras entre estes três grupos são fluídas, e que raramente dialogam entre si, apesar de partilharem o mesmo objecto de estudo.

5 (Disponível a partir de: https://hls.harvard.edu/today/are-you-an-ai-idealist-alarmist-or-skeptic/)

Falta-nos, sem terreno comum onde dialogar, o que, noutro momento e contexto, permitiu-nos conter um risco de escala semelhante. Em meados do século passado, quando a humanidade descobriu a capacidade de libertar a energia nuclear, foi preciso construir, em simultâneo, uma narrativa comum sobre o risco e os mecanismos para o conter: a Agência Internacional de Energia Atómica, em 57, e o Tratado de não proliferação de armas nucleares em 68, procuraram dar forma jurídica a essa "*confiança*" partilhada, submetendo um poder - sem precedentes - a limites, a verificação e a controlo. Hoje, agora em 2026, temos uma nova disputa em aberto. Teremos aptidão idêntica para perceber(mos) e a realidade que nos absorve e para confiar(mos) numa narrativa comum que nos permita reganhar controlo e confiança?

Com uma ironia que não perdeu a sua força, em 2003 Nick Bostrom[6] propôs o seguinte exercício mental: uma inteligência artificial encarregada de uma tarefa tão trivial como produzir clipes de papel poderia, se suficientemente competente e desprovida de quaisquer valores humanos, adequadamente representados nas suas finalidades e restrições, converter o planeta inteiro - e a humanidade com ele - em matéria-prima para mais clipes. A força desta imagem não está no exagero retórico. Está naquilo que ela exclui, propositadamente, da equação: os valores humanos. O maximizador de clipes, *per se*, não repudia ninguém. Simplesmente não vê pessoas onde vê átomos disponíveis para o seu fim. É a instrumentalização levada à sua forma mais pura. E é, com efeito, aquilo que a fórmula kantiana do fim em si («*a pessoa como fim em si mesma*») existe para impedir.

Mais de duas décadas depois, agora com os «*concierges de bolso*» por todo o lado - que em 2003 não existiam com a actual capacidade de redefinição da *nova normalidade digital* - os factos documentados acrescentam uma dimensão particularmente concreta à equação. Um dos mais recentes relatórios de inteligência sobre ameaças, publicado por uma empresa responsável por um dos *frontier labs* de inteligência artificial[7] do mundo ocidental, identificou dezenas de operações, desmanteladas, com tentativas de utilização desse sistema para actividades maliciosas, incluindo operações cibernéticas, de vigilância, desinformação, investigação biológica de risco e desenvolvimento de armamento convencional. Tendo sido

6 (Disponível em: https://nickbostrom.com/papers/ethical-issues-in-advanced-ai/)
7 (*Vide* https://www.anthropic.com/threat-intelligence-report-september-2026)

sinalizados como tentativas de utilização indevida dos sistemas, não podem porém, sem mais, ser lidos como prova de que tais objectivos tenham sido concretizados.

O risco, e o perigo, documentados não correspondem necessariamente ao cenário de uma máquina indiferente à sobrevivência humana, na prossecução, incondicional, de um objectivo absurdo. Vieram, precisamente, de pessoas, concretas, com intenções definidas, no uso de uma ferramenta cada vez mais competente para servir fins que, em abono da verdade, já eram potencialmente maliciosos muito antes de qualquer algoritmo existir. São problemas distintos, mas não incompatíveis. Bostrom ajudou-nos a compreender o perigo de um sistema cujo desempenho não incorporou adequadamente valores humanos. A realidade, em 2026, sublinha que o perigo mais imediato não é necessariamente o de uma máquina que decide *sozinha*. É o da pessoa que, através dela, decide - ou pretende(rá) decidir - por todos os outros.

A responsabilidade humana, porém, não se esgota na capacidade individual de reconhecer os riscos. Reclama, igualmente, um enquadramento jurídico que permita orientar, limitar e fiscalizar o desenvolvimento e a utilização destas tecnologias. É neste plano que se situa o Regulamento (UE) 2024/1689, o Regulamento Europeu da Inteligência Artificial (RIA, ou *EU AI Act*), entretanto alterado pelo Regulamento (UE) 2026/1744, no âmbito do pacote *Digital Omnibus on AI*.

Ao procurar estabelecer um quadro harmonizado, assente no risco, para a colocação no mercado, a colocação em serviço e para a utilização de sistemas de inteligência artificial, o RIA procura fixar uma gramática jurídica, comum, para uma tecnologia que atravessa sectores, instituições e práticas sociais muito diferentes. Todavia, a existência de uma *gramática normativa* não assegura, por si só, que as suas exigências sejam compreendidas, aplicadas e fiscalizadas, nem que se construa uma cultura comum de responsabilidade. A questão passa, por isso, pela concretização das opções de governação e execução que o próprio Regulamento deixa aos Estados-Membros, sem esquecer a aplicação directa das suas disposições e as obrigações que delas já decorrem.

É precisamente à concretização nacional desse quadro que se dedica uma proposta de anteprojecto, *de lege ferenda*, de uma lei de execução do Regulamento (UE) 2024/1689 (*EU AI Act*). Entre as soluções propostas, pontificam uma arquitectura nacional de governação, a atribuição à ANACOM das funções de ponto de contacto único e de autoridade de fiscalização do mercado, um Conselho de Coordenação e um Registo Nacional de Sistemas

de Inteligência Artificial. Com efeito, estará em causa, não apenas a definição de competências e procedimentos, mas procurar assegurar que o quadro europeu se concretiza num modelo institucional adequado às especificidades do ordenamento jurídico português. É neste intervalo entre o quadro europeu e as (futuras) escolhas nacionais de execução que a discussão jurídica se torna particularmente concreta, um pouco à semelhança do que já fizéramos, em 2019. Neste ponto a discussão aproxima-se daquela que já travámos a propósito da lei de execução nacional do Regulamento Geral sobre a protecção de dados pessoais (RGPD) cuja eficácia dependia, também aí, em parte, das opções institucionais e normativas tomadas no plano nacional.

A definição de um quadro jurídico para a inteligência artificial é, na verdade, mais um passo rumo às condições em que a confiança digital pode assentar. Mas a confiança não depende apenas das normas que enquadram o desenvolvimento e a utilização destas tecnologias, nem das instituições convocadas à sua aplicação. Depende também, e muito, da integridade dos próprios sistemas em que as decisões são tomadas e executadas.

Com efeito, e para que não resvalemos, *inconscientemente*, numa «*democracia dos algoritmos*», afastando a pessoa humana do seu centro, importa sublinhar que a confiança institucional pressupõe confiança em sistemas íntegros, onde as decisões correm - e correrão. A confidencialidade, a integridade, a autenticidade e a disponibilidade dos dados e das comunicações não são meros requisitos técnicos de uma sociedade da informação. São condições fundamentais da sua existência enquanto comunidade social, cultural, económica e política, encontrando respaldo em diversas garantias constitucionais. Reconhecer a sua importância implica, por tal, atravessar a linha de definição de garantias e assegurar mecanismos capazes de prevenir violações e de lhes dar resposta se e quando estas ocorram.

É neste ponto que cibersegurança e combate ao cibercrime acabam por confluir. As duas políticas protegem, no fundo, bens jurídicos que se cruzam, olhando-os de ângulos diferentes - um pela prevenção de vulnerabilidades e riscos, outro pela repressão de condutas ilícitas. Esta complementaridade torna relevante a articulação entre as medidas de identificação e prevenção de vulnerabilidades e a resposta jurídica aos actos que as exploram. Quem tem competência técnica para aceder sem autorização a um sistema detém, por essa mesma via, informação sobre vulnerabilidades de que a prevenção depende, seja ou não essa pessoa um cibercriminoso. Talvez por isso se perceba a proposta, tão (aparentemente) contra-intuitiva

quão necessária, de mecanismos legais que permitam, em condições estritamente delimitadas, afastar a punição de certos factos potencialmente susceptíveis de integrar crimes de acesso ou intercepção ilegítimos. O artigo 8.º-A da Lei do Cibercrime prevê essa possibilidade quando se encontrem cumulativamente preenchidos os pressupostos legais: a actuação deve ter por finalidade exclusiva a identificação de vulnerabilidades e contribuir para a segurança do ciberespaço, através de uma divulgação responsável; não pode visar a obtenção de vantagem económica, sem prejuízo da remuneração da actividade profissional; as vulnerabilidades identificadas devem ser comunicadas imediatamente após a actuação, ao proprietário ou à pessoa designada para gerir o sistema, produto ou serviço e, quando aplicável, ao titular dos dados obtidos; a actuação tem de ser proporcional e estritamente limitada ao necessário, sem perturbar o funcionamento dos sistemas, eliminar, deteriorar ou copiar dados informáticos sem autorização, ou produzir os efeitos prejudiciais proibidos pela norma; e não pode violar as regras em matéria de protecção de dados pessoais. O objectivo não é transformar o acesso ilegítimo numa actividade genericamente permitida, mas reconhecer que a protecção efectiva dos sistemas pode exigir uma resposta juridicamente diferenciada a condutas que, embora tecnicamente *semelhantes*, não são necessariamente equivalentes quanto às suas circunstâncias, finalidade e consequências.

Parece ser esta a lógica que o novo artigo 8.º-A da Lei do Cibercrime começa a acolher entre nós. A questão merece, porém, ser tratada com o rigor que o Direito Penal exige. A eventual não punibilidade depende do preenchimento dos pressupostos e limites definidos na lei, não apenas da declaração de uma intenção *benigna* ou da invocação de um interesse público. Cremos ser esta fronteira, que a lógica subjacente ao novo artigo procura delimitar, a que separa a competência técnica colocada ao serviço da segurança daquela que é movida pela agressão. Porém, a sua aplicação prática levanta questões que não se resolvem pela simples oposição entre o cibercriminoso e a "*ética*".

Na verdade, a questão aqui controvertida é, objectivamente, bem mais complexa. O mesmo comportamento técnico, nele incluindo um acesso não autorizado, pode assumir significados e consequências muito diferentes consoante as circunstâncias concretas em que ocorre. Seria excessivamente simplificador fazer depender essa distinção apenas da intenção do actor. Desde logo porque a intenção é uma realidade humana móvel, permeável, difícil de apreender e, sobretudo, incapaz de explicar por si só a totalidade de um comportamento. Alguém pode

aceder ilicitamente a um sistema para obter vantagem, exigir posteriormente um resgate e, depois de identificado, colaborar com as autoridades ou contribuir para a correcção da vulnerabilidade. O comportamento posterior pode ser relevante e produzir consequências próprias mas não transforma retroactivamente a natureza do comportamento e da intenção iniciais. Do mesmo modo, alguém que começa por identificar uma vulnerabilidade pode, perante uma oportunidade, adoptar uma conduta diferente daquela que inicialmente parecia desenhar-se.

É por isso que a distinção juridicamente relevante não pode assentar apenas na intenção declarada ou presumida. Exige olhar para o comportamento concreto, para a existência ou ausência de autorização, para as circunstâncias em que o acesso ocorreu, para os actos praticados sobre o sistema e para as consequências produzidas. A intenção pode ser relevante. As circunstâncias podem mudar. A oportunidade pode ser decisiva. O que não podemos fazer é transformar a intenção numa fronteira demasiado simples para uma realidade humana que, pela nossa própria natureza, o não é.

A criminalização do acesso e da intercepção ilegítimos tem uma genealogia internacional, que vai da Convenção de Budapeste à recente Convenção das Nações Unidas contra a Cibercriminalidade. Um regime como o do artigo 8.º-A, que delimita legalmente a não punibilidade de certas condutas associadas ao *hacking* ético e a sujeita a condições cumulativas, não tem de fragilizar o sistema penal: o acesso sem autorização pode estar associado a comportamentos muito diferentes, e é precisamente esta distinção, nos limites definidos pela lei, que importa reconhecer.

Porém, a necessidade de distinguir condutas não se coloca apenas perante o acesso técnico a um sistema. Coloca-se também perante a própria organização da criminalidade digital, quando as ferramentas e os serviços necessários para executar ataques passam a estar disponíveis a qualquer um, mesmo a quem não possui conhecimentos técnicos especializados. É esse o fenómeno do *Cybercrime-as-a-Service* (*CaaS*).

O *CaaS* constitui uma nova etapa na evolução do cibercrime. Ferramentas e serviços maliciosos são disponibilizados em mercados e fóruns da *dark web* e através de serviços de comunicações encriptadas, permitindo a utilizadores com poucas aptidões informáticas executar fraudes e ataques que, no passado, exigiam conhecimentos e meios muito mais

especializados. Ataques de negação de serviço, campanhas de *phishing*, *cryptojacking* e *ransomware* são exemplos de actividades que podem ser facilitadas por este mercado ilícito. As ligações deste fenómeno ao crime organizado e as suas possíveis intersecções com outras formas de ameaça, incluindo o terrorismo, a ciberguerra e a ciberespionagem, mostram que já não estamos apenas perante uma categoria residual do Direito Penal, mas ante uma infraestrutura de ameaça e risco que atravessa toda a vida colectiva. A utilização de criptomoedas, o alcance transnacional das plataformas e a diversidade de intervenientes tornam mais complexa a identificação dos responsáveis e a resposta jurídica. Importa, por isso, compreender não apenas as ferramentas e os serviços disponibilizados, mas também as ligações a outras formas de criminalidade, os prejuízos causados e as insuficiências que ainda possam existir na resposta do ordenamento jurídico português.

Depois da infraestrutura, e das condições de confiança possíveis que a sustentam, regressamos à pessoa que nela poderá habitar e construir-se. Em nenhum outro texto desta edição essa construção é tão literal como na vida extraordinária de Ibelin. Mats Steen, através do avatar que ergueu no universo de «*World of Warcraft*», não fugiu à sua existência física: prolongou-a e, nalguns planos, corrigiu-a ali, no ilogismo onde o corpo doente a limitava e a comunidade digital a expandia.

A leitura que o artigo propõe, cruzando biologia social e Direito, ajuda a demonstrar algo a que a nossa intuição jurídica nem sempre consegue responder adequadamente: uma segunda existência social, mediada digitalmente, pode, para quem nela vive, pesar tanto quanto, ou até mais do que, a existência social física? E podem as comunidades virtuais funcionar, para pessoas com deficiência ou em situações de exclusão estrutural, como palco de inclusão e reconhecimento que a vida física lhes nega/ou?

Reconhecer isto obriga a repensar o estatuto jurídico dessas comunidades sem as tratar como simulacro inferior de uma vida «*real*». Para quem nelas encontra aquilo que a exclusão lhes nega/ou, essa vida é, também ela, inteiramente real. A questão não é saber se a existência digital merece tutela por imitar a vida física, mas compreender o que nela se manifesta da própria pessoa, com toda a dignidade humana que lhe é inerente. Neste conflito, como deve o Direito ponderar? Deverá o Direito ponderar a partir daquilo que sempre o justificou: a pessoa concreta, como fim em si mesma, onde quer que ela se manifeste? E as plataformas?

Que responsabilidade lhes compete na preservação dos espaços em que essa vida, social, se constrói?

Se é a pessoa que decide através da máquina, é também na pessoa que começa a responsabilidade pelo modo como essa decisão é construída, implementada, executada e/ou utilizada. A confiança não pode, por isso, ser apenas *aquilo que se exige ao sistema*. É também aquilo que se constrói através da responsabilidade de quem o pensa, desenha, implementa e utiliza. E esta responsabilidade não se esgota numa fiscalização externa: constrói-se, também, dentro das próprias estruturas, quando estas aceitam transformar princípios abstractos em compromissos concretos e verificáveis. É nesse outro plano da confiança, o da responsabilidade assumida, que encontramos os Códigos de Conduta em matéria de protecção de dados. É nestes que essa responsabilidade assumida ganha forma *jurídica*.

Enquanto instrumentos de *soft law*, os códigos de conduta permitem que os responsáveis pelo tratamento e subcontratantes traduzam os princípios e as obrigações do Regulamento em compromissos concretos, sujeitos a mecanismos de acompanhamento e de controlo.

Não sendo figuras novas, porquanto a Directiva 95/46/CE já os previa, o RGPD conferiu-lhes um enquadramento mais habilitado. Onde antes havia um preceito próprio, encontramos agora os artigos 40.º e 41.º: o primeiro regula os códigos de conduta e a sua finalidade; o segundo estabelece o quadro da respectiva monitorização, sem prejuízo de outras disposições dispersas pelo Regulamento que lhes atribuem relevância e que os próprios artigos 40.º e 41.º não esgotam.

Um código de conduta não substitui a lei. Tradu-la para a linguagem concreta de um sector, tornando tangível aquilo que, na gramática do Regulamento, permanece por vezes difícil de aplicar no terreno. É este "*potencial de tradução*" que continua por explorar, sobretudo - como sublinha um dos artigos desta edição – nas Administrações Públicas. Com efeito, estas poderiam beneficiar de instrumentos sectoriais de concretização e responsabilização. Menos rígidos do que a lei, mais vinculativos do que uma simples boa prática, juridicamente enquadrados e sujeitos a mecanismos próprios de controlo, os Códigos poderiam, talvez, constituir um elo entre a Administração, que descrevemos atrás - a que se *procedimentaliza por algoritmo*, sem perder as suas garantias - e a confiança que só se sustenta quando as

regras deixam de ser abstractas e se tornam, também elas, práticas concretas e compreensivas para as pessoas e pelas pessoas.

O fio que aqui nos une - dignidade e confiança – atravessa, por fim, também a organização empresarial, que por vezes aparenta estranheza à linguagem dos direitos fundamentais. Numa fusão ou aquisição, os dados pessoais não são um activo a inventariar como se fossem existências de um qualquer armazém. O Regulamento Geral sobre a Protecção de Dados não prevê um regime específico para estas operações, e é essa ausência de regime específico que obriga a distinguir, com rigor, o risco próprio da transacção, os riscos que já existiam antes dela e os que podem surgir da integração numa nova estrutura. Até porque concluir um negócio não amplia, só por si, a licitude (ou ilicitude) do tratamento de dados que nele se transfere. Ademais, a omissão (de um regime autónomo, específico) não afasta as obrigações gerais relativas à licitude, à finalidade, à transparência, à minimização e à segurança dos tratamentos.

Para avaliar a prontidão do adquirente para essa integração, a metodologia proposta oferece um instrumento prático: permite determinar se, e em que condições, os dados da sociedade-alvo podem continuar a ser tratados, combinados com outros conjuntos ou migrados antes de ocorrerem alterações relevantes. A análise distingue os riscos da própria transacção, os riscos preexistentes e os riscos de integração, procurando assegurar que as questões jurídicas relevantes são apreciadas antes de mudanças materialmente significativas no tratamento. Reconhecer isto é reconhecer que a integração pós-aquisição é, também ela, uma questão de confiança. Sem olvidar, há sempre pessoas, dos dois lados do negócio, cuja vida os dados descrevem e cujos direitos e expectativas legítimas não se alteram só porque a organização empresarial mudou de proprietário.

É sempre assim: de pessoas, para pessoas, pelas pessoas.

Não como epígrafe que se declara e se esquece. Mas como verificação que se repete, com nomes e conceitos diferentes, em cada um dos textos que dão corpo à nova edição. Da sala de audiências ao balcão administrativo. Da perda silenciosa da capacidade de pensar à necessidade de saber ler o mundo que a tecnologia nos entrega, com espírito crítico e sem pressa, passando pela responsabilidade de quem cuida. Da gramática jurídica da inteligência artificial à vulnerabilidade de um sistema informático. Da criminalidade digital à

vulnerabilidade de quem nele deposita a sua vida digital. Da Administração Pública à sala de reuniões de uma fusão empresarial.

É sempre a mesma dignidade a ser testada.

E é sempre a mesma confiança, frágil e por reconstruir, que esta décima quarta edição convida a percorrer, ao longo de treze artigos científicos.

Num excurso para nos inquietar o pensamento.

Devagar.

Lentamente.

Como quem sabe que aquilo que vale a pena não se consome na fugacidade de dois toques num ecrã.

Faculdade de Direito da Universidade de Lisboa - Instituto de Investigação Interdisciplinar do Direito - IURIS

Centro Investigação Jurídica do Ciberespaço - CIJIC

Setembro de 2026

Nuno Teixeira Castro

PS: Feliz aniversário «*Cyberlaw-by-CIJIC*» pelo teu X aniversário. Parabéns a todos que fazem(os) parte desta viagem. Bem-haja a todos os que nos têm acompanhado e aos tantos outros que se nos juntarão. O que o verdadeiramente permanece e constrói uma comunidade é a relação, a conexão humana, entre quem cria, quem colabora, quem lê e quem se junta. O meu «Obrigado» a todos. Boas leituras e melhores discussões.



CYBERLAW

BY CIJIC

DOUTRINA



INTELIGÊNCIA ARTIFICIAL E ATIVIDADE JURISDICIONAL: ENTRE A EPISTEMOLOGIA JURÍDICA E OS DESAFIOS DE GOVERNAÇÃO

NUNO COELHO *

* Juiz Conselheiro do Supremo Tribunal de Justiça

RESUMO:

Partindo da integração acelerada da inteligência artificial na administração da justiça, este artigo situa a questão no plano da epistemologia jurídica, recusando tratar a tecnologia como um dado neutro. Propõe uma tipologia de três níveis de intervenção tecnológica na atividade jurisdicional - gestão, apoio à decisão e reflexão crítica - e equaciona o modelo IRAC com o esquema clássico de realização do direito, mostrando que a fronteira entre o que a inteligência artificial pode assistir e o que permanece exclusivo do juiz corresponde à distinção entre operações lógico-classificatórias e juízo tipológico. Projeta depois estas implicações na formação judiciária e na governação dos tribunais, concluindo que a pluralidade de níveis e modelos não dissolve, antes reconstitui, a unidade da função de julgar: não uma justiça substituída, mas uma justiça aumentada, sob responsabilidade do juiz e das instituições que o garantem.

Palavras-chave: Inteligência artificial; Atividade jurisdicional; Epistemologia jurídica; Hermenêutica jurídica; Governação judiciária; Formação judiciária.

ABSTRACT:

Starting from the accelerated integration of artificial intelligence into the administration of justice, this article situates the issue within the realm of legal epistemology, rejecting the treatment of technology as a neutral given. It proposes a three-tier typology of technological intervention in judicial activity - management, decision support, and critical reflection - and compares the IRAC method with the classic model of legal reasoning, demonstrating that the boundary between what artificial intelligence can assist and what remains exclusive to the judge corresponds to the distinction between logical-classificatory operations and typological judgment. It subsequently projects these implications onto judicial training and court governance, concluding that the plurality of levels and models does not dissolve, but rather reconstitutes, the unity of the judicial function. Not a replaced justice, but an augmented justice, under the responsibility of the judge and the institutions that guarantee it.

Keywords: Artificial intelligence; Judicial activity; Legal epistemology; Legal hermeneutics; Judicial governance; Judicial training.

"A técnica não é apenas um conjunto de meios, mas um modo de conhecimento, um pensamento em ato."
- Gilbert Simondon, Du mode d'existence des objets techniques, 1958
"O que distingue a civilização moderna não é a máquina em si, mas a sua aplicação sistemática ao pensamento, ao trabalho e à vida social."
- Lewis Mumford, Technics and Civilization, 1934
"O poder funda-se numa arquitetura epistémica: as tecnologias não apenas organizam o saber, mas moldam as formas como esse saber pode ser exercido e legitimado."
- Peter Kahl, The Epistemic Architecture of Power, 2025

I. INTRODUÇÃO. O EU MÚLTIPLO, A PROPÓSITO DE UMA EXPOSIÇÃO TEMPORÁRIA

Visitámos, no Museu de Arte Contemporânea Armando Martins, a exposição «O Eu como Múltiplo», terceira mostra temporária da Coleção MACAM, com curadoria de Carolina Quintela.[1] Um dos núcleos da exposição, centrado na obra de Paula Rego, apresenta a mente humana como uma instância que não opera isoladamente: entrelaça-se em fluxos de memória e sofre influências externas da história e da cultura na formação, sempre inacabada, de um subconsciente coletivo. Um segundo painel, o que dá título à mostra, define a identidade não como essência estática, mas como movimento - um diálogo contínuo entre o visível e o invisível, entre o ser e o tornar-se. E uma terceira peça, em néon, evoca uma história incorporada, uma manutenção perpétua do passado que habita os nossos contextos sociais como se de uma arrecadação se tratasse.

Estes três registos - a psique que se forma por sedimentação histórica e cultural, a identidade como processo e não como dado adquirido, a memória que persiste materialmente nos artefactos e nas instituições - oferecem, na sua justaposição, uma chave de leitura pouco comum, mas fecunda, para o problema que este texto pretende tratar: o modo como a

1Quintela, C. (2025). O eu como múltiplo [Catálogo de exposição]. Lisboa: MACAM - Museu de Arte Contemporânea Armando Martins.

atividade jurisdicional, enquanto prática de conhecimento e de decisão, se relaciona com sistemas de inteligência artificial que hoje penetram, com intensidade e velocidade distintas, no exercício da função de julgar. Se a identidade do sujeito - e, com ela, a do juiz enquanto sujeito epistémico - se constrói por sedimentação de gestos, memórias e influências, importa perguntar o que sucede quando parte desses gestos, dessas memórias de trabalho e desses processos de triagem e de síntese documental passam a ser mediados, ou mesmo pré-formatados, por sistemas artificiais de processamento de informação e de geração de texto.

É esta a questão que as três citações que encimam o texto ajudam a situar. Simondon via na técnica não um conjunto neutro de meios, mas um modo de conhecimento - um pensamento em ato.[2] Mumford, décadas antes de existir qualquer forma de inteligência artificial no sentido em que hoje a entendemos, sustentava que a civilização moderna se distingue não pela máquina em si, mas pela sua aplicação sistemática ao pensamento, ao trabalho e à vida social.[3] Kahl, já no nosso tempo, descreve o poder como algo que se funda numa arquitetura epistémica, na qual as tecnologias não só organizam o saber, como moldam as formas pelas quais esse saber é exercido e legitimado.[4]

Estas três posições, distantes entre si quase um século e enraizadas em tradições intelectuais diferentes - a filosofia da técnica, a história da tecnologia, a teoria política do conhecimento - convergem num ponto que interessa diretamente à reflexão jurídica: a tecnologia não é um instrumento exterior ao conhecimento e à autoridade que dele deriva, mas uma condição da sua própria formação e legitimação. Aplicada à administração da justiça, esta convergência obriga a repensar aquilo que tradicionalmente se considerou pertencer, sem mediação relevante, à competência exclusiva do juiz: o conhecimento dos factos, a interpretação da norma e a formação da decisão.

A administração da justiça, pilar fundamental de qualquer sociedade organizada, encontra-se hoje no limiar de uma transformação sem precedentes, impulsionada pela digitalização e pela integração da inteligência artificial.[5] Neste contexto, a IA é cada vez mais vista como parte integrante da transformação digital em geral, e não como um avanço tecnológico separado e

2Simondon, G. (1958). Du mode d'existence des objets techniques. Paris: Aubier-Montaigne.
3Mumford, L. (1934). Technics and civilization. Nova Iorque: Harcourt, Brace and Company.
4Kahl, P. (2025). The epistemic architecture of power: How knowledge control sustains authority in social structures (2.ª ed.). Londres: Lex et Ratio Ltd..
5Coelho, Nuno (2024). Organização do sistema judicial e jurisdição: Para uma sistemática alargada da realização do direito pelo juiz. Coimbra: Almedina, pp. 339-370.

isolado. Esta mudança transcende meras atualizações operacionais: redefine, de forma fundamental, o modo como a justiça é administrada, acedida e percecionada a nível global. A imperatividade desta metamorfose é sublinhada pela necessidade premente de maior eficiência, acessibilidade e transparência nos sistemas judiciais, sem perda das qualidades essenciais que marcam os seus fundamentos constitucionais.[6]

É neste quadro que se situa o objeto deste texto: a relação entre a tecnologia - designadamente os novos sistemas de inteligência artificial - e a atividade do juiz e dos tribunais. Interessa-nos, em particular, o impacto destes sistemas no domínio jurídico, nas profissões da justiça e, sobretudo, na atividade jurisdicional propriamente dita, entendida como o núcleo da função de julgar.

Esta relação pode ser observada a partir de diferentes dimensões, que não se excluem mas antes se implicam mutuamente:

a) do conhecimento e da formação, que interroga o modo como os juízes e os demais operadores da justiça adquirem, mantêm e atualizam a competência técnica exigida pelo exercício da função num ambiente já parcialmente mediado por sistemas artificiais;

b) da norma e da atividade interpretativa e aplicativa do direito, que questiona os efeitos da inteligência artificial sobre a interpretação e a aplicação do direito, sem as quais não há decisão jurisdicional que mereça esse nome;

c) da regulação ética e legal, atenta ao quadro normativo - europeu e nacional - que disciplina o uso destes sistemas no exercício de poderes públicos;

d) da organização e da gestão da atividade judicial, que trata a inteligência artificial como instrumento, e por vezes como fator de perturbação, da organização judiciária e da gestão processual; e

e) do processo e dos procedimentos, atenta às implicações da inteligência artificial sobre o processo e sobre as garantias que o estruturam.

Cada uma destas dimensões será retomada, com desenvolvimento próprio, nos pontos seguintes. O segundo ponto ocupa-se do enquadramento epistemológico da relação entre o social e o tecnológico, situando a questão jurídica num plano mais amplo de epistemologia geral, sem o qual a tecnologia correria o risco de ser tratada como um dado neutro. O terceiro

6Coelho, Nuno (2024). Ob. cit., pp. 339-370.

ponto avança para a relação entre a inteligência artificial e a atividade jurisdicional, identificando os modos como estes sistemas já intervêm, ou podem vir a intervir, na função de julgar. O quarto ponto propõe uma tipologia dos níveis de intervenção tecnológica na atividade jurisdicional, distinguindo graus de mediação que vão do auxílio instrumental à eventual substituição de segmentos do processo decisório. O quinto ponto detém-se na dimensão hermenêutica, discutindo o modelo IRAC - *Issue, Rule, Application, Conclusion* - e os limites que a tecnologia atual, e a previsível num futuro próximo, enfrenta perante a exigência de fundamentação própria da decisão judicial. O sexto ponto extrai as implicações formativas desta análise, esboçando os contornos de uma cultura epistemológica tecnológica que a formação judiciária terá de incorporar. O sétimo e último ponto fecha o texto com uma reflexão sobre os desafios de governação, coordenação e cooperação que a integração da inteligência artificial na justiça coloca às instituições, aos Estados e às organizações internacionais e profissionais que a eles se associam.

Se a exposição do MACAM sugere que o eu não é uma essência fixa, mas um processo contínuo de composição - de gestos, memórias e influências que se cruzam sem nunca se estabilizarem de vez - não parece de todo descabido perguntar se algo de semelhante não está a suceder à própria função de julgar.[7] Não se trata de substituir o juiz por uma máquina, mas de reconhecer a recomposição, sob pressão tecnológica, dos processos de conhecimento, de memória de trabalho e de decisão que até agora se consideravam pertencer, sem mediação relevante, ao foro exclusivo da consciência do julgador. É esta recomposição, nas suas várias dimensões e nos seus limites, que as páginas seguintes procuram descrever.

II. O SOCIAL E O TECNOLÓGICO. EPISTEMOLOGIA JURÍDICA

A introdução deste texto partiu de três citações - Simondon, Mumford, Kahl - que, na sua diversidade, convergem numa mesma tese: a técnica não é um meio neutro ao serviço de fins que lhe são exteriores, mas uma condição da própria formação e legitimação do conhecimento. Este segundo ponto retoma essa tese e procura dar-lhe densidade epistemológica, situando a relação entre o social e o tecnológico num plano que é,

7Nota metodológica: a aproximação entre o discurso curatorial da exposição e a reflexão jurídica que aqui se propõe constitui um recurso metodológico de abertura, e não uma tese sobre a natureza da função jurisdicional, cuja autonomia normativa e institucional não está, evidentemente, em causa.

simultaneamente, filosófico, sistémico e jurídico. Sem esta densificação, o que se dirá adiante sobre a inteligência artificial e a atividade jurisdicional correria o risco de permanecer descritivo, sem alcançar as razões profundas pelas quais a tecnologia interpela hoje, com tal intensidade, os fundamentos do próprio direito.

A génese cibernética da epistemologia tecnológica

A relação contemporânea entre o social e o tecnológico tem uma genealogia identificável, e essa genealogia chama-se cibernética. Norbert Wiener, ao fundar esta disciplina em meados do século XX, formulou um modelo geral de regulação assente em malhas de retroalimentação (*feedback loops*) e em mecanismos de controlo que, projetados inicialmente para sistemas mecânicos e biológicos, vieram a fornecer o vocabulário conceptual com que hoje se pensam tanto os sistemas artificiais como, por via da teoria dos sistemas sociais, os próprios sistemas sociais e jurídicos.[8] A cibernética ultrapassa, por isso, o estatuto de uma técnica entre outras: constitui uma epistemologia - um modo de conhecer fundado no controlo, na redução da contingência e na retroação - que se difundiu muito para além do laboratório de onde saiu.

Martin Heidegger via nesta difusão algo mais do que um progresso instrumental: para o filósofo alemão, a essência da técnica moderna não é ela própria algo técnico, mas um modo de desvelamento (*Entbergen*) que reduz o mundo a uma reserva permanente de recursos disponíveis para requisição — o que designou por *Gestell*, ou dispositivo de imposição.[9] Yuk Hui, no seu recente *Kant machine: Critical philosophy after AI*, retoma esta tese heideggeriana para a confrontar com a inteligência artificial contemporânea, sustentando que os modelos generativos atuais, por mais sofisticados que sejam, permanecem no âmbito do cálculo e da otimização recursiva, carecendo da capacidade de autodeterminação que Kant reservava à razão prática.[10] Se a epistemologia jurídica quer compreender o que está em jogo na integração da inteligência artificial na justiça, não pode ignorar esta genealogia: a questão

8Wiener, N. (1961). Cybernetics: Or control and communication in the animal and the machine (2.ª ed.). Cambridge, MA: MIT Press.
9Heidegger, M. (1977). The question concerning technology and other essays (W. Lovitt, Trad.). Nova Iorque: Garland Publishing. (Ensaio original, "Die Frage nach der Technik", 1954.)
10Hui, Y. (2026). Kant machine: Critical philosophy after AI. Londres: Bloomsbury Academic.

não se esgota em saber o que a tecnologia pode fazer; importa, sobretudo, saber que forma de conhecimento ela pressupõe e impõe.

Duas máquinas, duas racionalidades

A história da epistemologia ocidental oferece, a este propósito, uma chave de leitura para a inteligência artificial contemporânea que Yuk Hui sistematiza através da distinção entre duas arquiteturas cognitivas: a máquina cartesiana e a máquina *humeana*. A primeira corporiza a tradição racionalista: a inteligência é a execução mecânica de regras lógicas *a priori* sobre símbolos abstratos, tal como a jurisprudência dos conceitos concebeu, no século XIX, o juiz como aplicador silogístico de um sistema fechado de normas. A segunda corporiza a tradição empirista de Hume: a inteligência resulta da associação estatística de dados segundo relações de contiguidade e de frequência, sem qualquer necessidade lógica *a priori* - o paradigma que, formalizado por Ray Solomonoff na sua teoria da inferência indutiva, sustenta hoje as redes neuronais e a aprendizagem automática, e que no domínio judiciário se projeta na chamada justiça preditiva.[11]

Nenhuma das duas máquinas, por si só, dá conta do que se exige ao juízo jurídico. Foi para superar o impasse entre o dogmatismo cartesiano e o ceticismo *humeano* que Kant formulou a distinção, decisiva para o que se dirá no quinto ponto deste texto, entre o juízo determinante (*Bestimmendes Urteil*) e o juízo reflexionante (*Reflektierendes Urteil*): enquanto o primeiro subsume um particular sob uma regra já dada, o segundo parte do particular, sem regra prévia disponível, e ascende heuristicamente até um universal que tem de descobrir.[12] A decisão judicial, na medida em que responde a casos concretos e irrepetíveis, aproxima-se estruturalmente deste segundo tipo de juízo - e é precisamente aqui que a promessa e o limite da inteligência artificial se separam com maior nitidez: um sistema que opera por subsunção determinante ou por indução estatística não está, pela sua própria arquitetura, talhado para o juízo reflexionante que a singularidade do caso jurídico reclama.

Direito como sistema epistémico autónomo: autopoiese, alopoiese e jurisprudencialismo

11Solomonoff, R. (1964). A formal theory of inductive inference (Parts I & II). Information and Control, 7(1–2), 1–22, 224–254.

12Kant, I. (2007). Critique of judgment (J. C. Meredith & N. Walker, Trad.). Oxford: Oxford University Press.

Se a cibernética forneceu o vocabulário do controlo e da retroação, foi Niklas Luhmann quem o transpôs, de modo mais influente, para a teoria do direito. Na sua Teoria dos Sistemas Sociais, o sistema jurídico é concebido como um subsistema autopoiético, que produz e reproduz autonomamente os seus próprios elementos - decisões, normas, qualificações - a partir da sua rede interna de comunicações, operando através do código binário lícito/ilícito (*Recht/Unrecht*).[13] Este fechamento operacional não isola o direito do seu ambiente social; é, pelo contrário, a condição da sua abertura cognitiva, que se realiza através de mecanismos de acoplamento estrutural - a Constituição, o contrato - pelos quais o sistema jurídico absorve a complexidade exterior e a traduz para a linguagem do seu próprio código.

Esta arquitetura sistémica ilumina, por contraste, o risco que a integração acrítica da tecnologia no exercício da função judicial comporta: o risco de alopoiese, ou seja, a quebra da autonomia operacional do direito quando a produção das decisões passa a ser determinada, direta ou indiretamente, pelos códigos e pelas exigências funcionais de outros sistemas - a eficiência económica, a produtividade quantitativa, a conformidade cega com um resultado algorítmico.[14] Quando isso sucede, o direito não deixa de existir enquanto conjunto de textos, mas perde a densidade normativa que o distingue de uma mera técnica de gestão comportamental.

É por esta via sistémica, mas também por uma via diretamente jurisprudencial, que a metodologia jurídica portuguesa contemporânea recusa a redução do juízo judicativo a uma tecnologia de decisão. Para António Castanheira Neves, a metodologia jurídica não é um catálogo de procedimentos técnicos, mas uma reflexão crítica sobre a própria realização do direito: o caso concreto, e não a norma abstrata, constitui o *prius* metodológico do pensamento jurídico, e a interpretação e a aplicação não são fases sucessivas, mas um processo unitário em que facto e norma se co-determinam no círculo hermenêutico.[15] Nesta linha, também João Baptista Machado situou o direito não na esfera da natureza inorgânica nem na mera biosfera, mas naquilo a que chamou a noosfera - a dimensão cultural e normativa da existência humana - sustentando que a atividade judiciária atualiza permanentemente um discurso legitimador cuja validade repousa na adesão comunitária a

13Luhmann, N. (2004). Law as a social system (K. A. Ziegert, Trad.). Oxford: Oxford University Press.
14Neves, M. (1994). A Constitucionalização Simbólica. São Paulo: Editora Acadêmica.
15Castanheira Neves, A. (1993). Lições de metodologia jurídica. Coimbra: Coimbra Editora.

valores de justiça, e não numa técnica neutra de resolução de conflitos.[16] Entre a arquitetura sistémica de Luhmann e a reconstituição jurisprudencialista de Castanheira Neves há tensões que este texto não pretende resolver; mas ambas convergem, por caminhos diferentes, na mesma conclusão negativa: o direito não é redutível a informação processada segundo um código que lhe é alheio.

A cognição em rede: sobreposições institucionais e a delegação de agência aos códigos

Esta relação entre o sistema cognitivo e o mundo tecnológico foi já objeto de desenvolvimento próprio noutro lugar, a que convém agora regressar com maior densidade.[17] Sustentou-se aí que o sistema cognitivo está imerso no mundo social e tecnológico, com ele interage e com ele se potencia - uma tese que, para além da extensão da mente já enunciada, ganha densidade adicional com os contributos de Paul Smart, que reviu criticamente as controvérsias em torno da cognição estendida, e de Giulio Ongaro, Doug Hardman e Ivan Deschenaux, para quem a mente estendida não constitui uma singularidade excêntrica da filosofia da cognição, mas antes um traço central e recorrente da própria atividade cognitiva.[18]

As alterações paradigmáticas na organização social causadas pelas mutações nas redes de comunicação e interação inerentes à revolução tecnológica em curso não deixam, por isso, de afetar o próprio direito. A operacionalidade deste é questionada pela proliferação de padrões em rede e pela intervenção de quasi-atores, mais ou menos complexos e multinível, e de plataformas, em que as figuras da conexão e da parceria contratual assumem relevância crescente na fronteira entre a regulação privada e a pública - uma linha de investigação que Karl-Heinz Ladeur situou na epistemologia social da decisão e na transição da subsunção do caso para a ponderação e para uma semântica de redes, que Gunther Teubner desenvolveu a propósito da matriz anónima de violações de direitos humanos por atores transnacionais

16Baptista Machado, J. (2017). Introdução ao direito e ao discurso legitimador. Coimbra: Almedina. (1.ª ed. 1983.)

17Coelho, Nuno (2024). Ob. cit., pp. 339-343.

18Clark, A., & Chalmers, D. (1998). The extended mind. Analysis, 58(1), 7–19; Smart, P. (2017). Extended cognition and the internet: A review of current issues and controversies. Philosophy & Technology, 30, 357–390; Ongaro, G., Hardman, D., & Deschenaux, I. (2022). Why the extended mind is nothing special but is central. Phenomenology and the Cognitive Sciences.

privados e das redes como contratos conectados, que Thomas Vesting analisou como um problema de autonomia do direito face à formação de padrões de rede, e que François Ost e Michel van de Kerchove sintetizaram na fórmula de uma passagem da pirâmide à rede, em busca de uma teoria dialética do direito.[19] Esta mesma linha de reflexão remete para uma epistemologia histórica e social e para um contexto prático-institucional em que o conhecimento dos fenómenos coletivos e da rede vem produzindo um novo paradigma da cultura jurídica e científica, como o próprio Vesting viria a aprofundar a propósito da crise do conhecimento comum e dos meios do direito.[20]

Esta dinâmica institucional pode ser lida, num registo mais propriamente sociológico e organizacional, através da distinção entre as componentes institucional, estrutural e tecnológica da dinâmica organizacional proposta por Bruno Maggi, do acoplamento estrutural entre as realizações técnicas e o sistema social por via da redução da complexidade - retomando aqui, num registo complementar ao da autopoiese já referida, a teoria sociológica do risco de Luhmann - e da ordem da tecnologia enquanto complexidade e controlo num mundo conectado que Jannis Kallinikos descreveu, mais tarde desenvolvida na sua proposta de uma governação através da tecnologia.[21] As sobreposições institucionais que destas mutações resultam podem ser caracterizadas com recurso às noções de ator, de rede e de agência: nelas ocorrem processos de transferência das regulações normativas, e dos respetivos comandos, para os códigos informáticos, numa estética de sobreposição e

19Ladeur, K.-H. (2007). The role of contracts and networks in public governance: The importance of the "social epistemology" of decision making. Indiana Journal of Global Legal Studies, 14(2), 329–351; Ladeur, K.-H. (2017). The transformation of legal semantics in postmodern society. Revista de Informação Legislativa, 54(215), 13–26; Teubner, G. (2006). The anonymous matrix: Human rights violations by "private" transnational actors. Modern Law Review, 69(3), 327–346; Teubner, G., & Collins, H. (2011). Networks as connected contracts. Oxford/Portland: Hart Publishing; Vesting, T. (2004). The autonomy of law and the formation of network standards. German Law Journal, 5(6), 639–668; Ost, F., & van de Kerchove, M. (2002). De la pyramide au réseau? Pour une théorie dialectique du droit. Bruxelas: Publications des Facultés universitaires Saint-Louis.

20Vesting, T. (2017). Modern law and the crisis of common knowledge. Revista de Informação Legislativa, 54(214), 11–29; Vesting, T. (2018). Legal Theory and the Media of Law. Cheltenham/Northampton: Edward Elgar Publishing.

21Maggi, B. (1984). Teoria dell'organizzazione e sociologia del diritto. Em U. Scarpelli & V. Tomeo (Eds.), Società, norme e valori: Studi in onore di Renato Treves (pp. 285–307). Milão: Giuffré; Luhmann, N. (1993a). Risk: A sociological theory. Berlim/Nova Iorque: Walter de Gruyter; Kallinikos, J. (2005). The order of technology: Complexity and control in a connected world. LSE Research Online; Kallinikos, J. (2011). Governing through technology: Information artefacts and social practice. Hampshire/Nova Iorque: Palgrave Macmillan.

acumulação de materiais que Giovan Francesco Lanzara designou, tomando de empréstimo um termo às artes, por assemblages.[22]

Por via deste fenómeno de agência, ou delegação, os componentes processuais definidos pela lei são inscritos na própria tecnologia. A delegação para as máquinas é dupla: as máquinas realizam ações anteriormente desempenhadas por humanos - automação - ao mesmo tempo que orientam os humanos na execução das suas tarefas, sendo que, quanto mais a justiça eletrónica se difunde, mais a autoridade da lei processual se delega para os códigos de software, como documentaram Lanzara e, mais recentemente, Francesco Contini, a propósito da transformação das interações entre humanos, direito e tecnologia nos processos judiciais mediados por inteligência artificial.[23]

É este o terreno em que se pode legitimamente perguntar se poderemos vir a falar de uma "*artificial judicial decision making*" - interrogação que já em 2001 Cass Sunstein formulava a propósito da inteligência artificial e do raciocínio jurídico, que Alexandre Libório Dias Pereira desenvolveu sob a designação de *lex informatica*, *ius ex machina* e justiça artificial, e que Ignacio Cofone retomou a propósito da tomada de decisão judicial assistida por inteligência artificial, distinguindo sempre a padronização obtida pela análise mecânica e generativa dos dados e a analogia jurídica, que é sempre realizada com apelo a princípios e valores.[24] Ricardo Campos sintetiza este conjunto de mutações na fórmula de metamorfoses do direito global, resultantes da interação entre direito, tempo e tecnologia, que reorientam os próprios processos de produção do direito entre o Estado-nação, as organizações transnacionais e as plataformas digitais autorreguladas.[25] A hibridização do direito, nestes

22Lanzara, G. F. (2009). Building digital institutions: ICT and the rise of assemblages in government. Em F. Contini & G. F. Lanzara (Eds.), ICT and innovation in the public sector (pp. 9–48). Nova Iorque: Palgrave MacMillan.

23Lanzara, G. F. (2009). Ob. cit., pp. 9-13; Lanzara, G. F. (2014). The circulation of agency in judicial proceedings: Designing for interoperability and complexity. Em F. Contini & G. F. Lanzara (Eds.), The circulation of agency in e-justice. Dordrecht/Heidelberga/Nova Iorque/Londres: Springer; Contini, F. (2020). Artificial intelligence and the transformation of humans, law and technology interactions in judicial proceedings. Law, Technology and Humans, 2(1), 4–18.

24Pereira, A. L. D. (2008). Lex informática, ius ex machina e justiça artificial. Em J. de Figueiredo Dias, J. J. Gomes Canotilho & J. de Faria Costa (Orgs.), Ars Iudicandi — Estudos em Homenagem ao Prof. Doutor António Castanheira Neves, Vol. I (pp. 817–886). Coimbra: Coimbra Editora; Sunstein, C. R. (2001). Of artificial intelligence and legal reasoning. Public Law & Legal Theory, University of Chicago Law School, Working Paper n.º 18; Cofone, I. (2021). AI and judicial decision-making. Em F. Martin-Bariteau & T. Scassa (Eds.), Artificial intelligence and the law in Canada. Toronto: LexisNexis.

25Campos, Ricardo (2022). Metamorfoses do direito global: Sobre a interação entre direito, tempo e tecnologia. São Paulo: Contracorrente.

termos, não é uma hipótese distante: é a descrição mais rigorosa do que já está em curso - e é o que a psicologia cognitiva experimental vem, como se verá de seguida, confirmar com dados de comportamento.

Sistema 3, rendição cognitiva e epistemia: a confirmação empírica de um risco epistemológico

O que a filosofia da técnica, a teoria do direito e a sociologia das organizações descrevem em termos sistémicos e institucionais encontra hoje uma confirmação inesperada na psicologia cognitiva experimental. Steven Shaw e Gideon Nave, investigadores da Wharton School, deram à tese da mente estendida, já enunciada, um tratamento experimental rigoroso através da Teoria Tri-Sistema da Cognição, que introduz, a par do Sistema 1 (intuitivo) e do Sistema 2 (deliberativo) da literatura clássica sobre o raciocínio humano, um terceiro sistema - o Sistema 3, cognição artificial externa ao cérebro. Em três experiências pré-registadas com mais de mil participantes, os autores demonstraram um fenómeno a que chamaram rendição cognitiva (*cognitive surrender*): quando um sistema de inteligência artificial está disponível, os participantes consultam-no na maioria das vezes e seguem as suas respostas mesmo quando estas são erradas, numa proporção próxima de quatro em cada cinco casos, fazendo-o com uma confiança que não diminui significativamente perante o erro. Apenas incentivos de desempenho associados a retroação imediata, item a item, atenuam parcialmente este padrão, sem, contudo, o eliminarem.[26]

Um grupo de investigadores da *Sapienza*, liderado por Walter Quattrociocchi, oferece o correlato epistemológico deste achado comportamental através do conceito de epistemia: uma condição estrutural em que a plausibilidade linguística substitui a avaliação epistémica, produzindo a sensação de saber sem que se tenha percorrido o trabalho de justificação que normalmente a fundamenta. Os mesmos autores identificam sete linhas de fratura entre a cognição humana e a artificial - ancoragem percetiva (*grounding*), segmentação da experiência (*parsing*), continuidade biográfica, motivação, raciocínio causal, metacognição e sensibilidade a valores - que, tomadas em conjunto, explicam por que motivo um sistema que apenas prevê estatisticamente sequências de símbolos pode parecer raciocinar sem, na

26Shaw, S. D., & Nave, G. (2026). Thinking-fast, slow, and artificial: How AI is reshaping human reasoning and the rise of cognitive surrender (working paper). Filadélfia: The Wharton School, University of Pennsylvania.

verdade, o fazer.[27] Não é difícil reconhecer, nestas sete fraturas, exatamente as dimensões que a reflexão jurisprudencialista, na secção anterior, reivindicava como próprias e insubstituíveis do juízo judicativo: a valoração, a intersubjetividade, a responsabilidade. A rendição cognitiva e a epistemia não são, portanto, curiosidades da psicologia experimental; são a tradução comportamental, medida em laboratório, do mesmo risco que Luhmann descrevia como alopoiese e que Castanheira Neves recusava sob a designação de tecnologia de decisão.

Da epistemologia à questão democrática

Esta análise não pode fechar-se sobre si mesma, num plano puramente epistemológico, sem tocar as suas implicações institucionais - que os pontos seguintes deste texto, sobretudo o sétimo, desenvolverão com mais vagar. David Gunkel formulou o que chamou a questão da máquina: em que medida sistemas autónomos podem reivindicar, ou ser objeto de, responsabilidades morais e jurídicas que até agora se pressupunham exclusivamente humanas.[28] Mark Coeckelbergh, situando o problema no plano da teoria democrática, mostrou como a integração da inteligência artificial na formação da opinião e do conhecimento cívico arrisca diminuir a agência epistémica dos cidadãos, ao substituir o exercício autónomo do juízo por bolhas epistémicas e por uma deferência acrítica ao conhecimento estatístico.[29] O paralelo com a atividade judiciária é direto: se a agência epistémica do cidadão é condição da democracia, a agência epistémica do juiz é condição da própria jurisdição, e o risco de rendição cognitiva descrito na secção anterior não é menos grave quando quem se rende é o julgador.

Claude Lefort ensinou que a democracia moderna se distingue por manter vazio o lugar do poder (*le lieu vide du pouvoir*), impedindo que qualquer instância - um partido, uma classe, ou hoje, acrescente-se, um sistema técnico - o ocupe de modo definitivo e incontestável.[30] Delegar em sistemas automatizados juízos que pressupõem interpretação e responsabilidade

27Quattrociocchi, W., Capraro, V., & Perc, M. (2025). Epistemological fault lines between human and artificial intelligence. arXiv:2512.19466.

28Gunkel, D. J. (2012). The machine question: Critical perspectives on AI, robots, and ethics. Cambridge, MA: MIT Press.

29Coeckelbergh, M. (2022). Democracy, epistemic agency, and AI: Political epistemology in times of artificial intelligence. AI and Ethics. https://doi.org/10.1007/s43681-022-00239-4

30Lefort, C. (1988). Democracy and political theory (D. Macey, Trad.). Cambridge: Polity Press.

arrisca preencher esse lugar vazio com uma autoridade que já não se deixa contestar, precisamente porque se apresenta como cálculo e não como decisão. Não se trata, com isto, de propor uma única resposta epistemológica fechada: correntes contemporâneas da teoria do direito, como a leitura rizomática que Anne Wagner e Sarah Marusek têm vindo a desenvolver, sugerem que a própria pluralidade e descentração do pensamento jurídico podem constituir recursos perante a complexidade tecnológica do presente, e não só riscos.[31] O que este ponto reteve, e que os pontos seguintes desenvolverão, é que a integração da inteligência artificial na atividade jurisdicional não é uma questão meramente instrumental de eficiência processual: é uma questão epistemológica de primeira ordem, que decide se o direito continuará a ser um sistema de conhecimento autónomo, fundado na responsabilidade e na valoração, ou se se deixará capturar por uma racionalidade que lhe é estranha. É com esta questão em aberto que se avança, no ponto seguinte, para a relação concreta entre a inteligência artificial e a atividade jurisdicional.

III. A IA E A ATIVIDADE JURISDICIONAL

O ponto anterior fechou com uma pergunta em aberto: se a integração da inteligência artificial na justiça é uma questão epistemológica de primeira ordem, o que significa isso, em concreto, para a atividade jurisdicional? Este terceiro ponto retoma essa pergunta e desce ao plano genérico da relação entre a inteligência artificial e a atividade do juiz e dos tribunais, preparando o quarto ponto, onde essa relação será decomposta nos seus diferentes níveis de intervenção tecnológica. Antes de perguntar o que a inteligência artificial pode fazer pela justiça, importa, contudo, perguntar o que é, afinal, a atividade jurisdicional que aqui está em causa.

O que é, afinal, a atividade jurisdicional?

Um relatório temático de 2013 do Institut des Hautes Études sur la Justice, da autoria de Antoine Garapon, Sylvie Perdriolle, Boris Bernabé e Charles Kadri, alertava já para o ponto de vista estreito de que partem as abordagens jurídicas clássicas ao apreciarem a função

31Wagner, A., & Marusek, S. (2026). Reimagining law through the rhizome - Fluidity, fracture, and forward movement [Editorial]. Undecidabilities and Law: The Coimbra Journal for Legal Studies, (6), 7-19. https://doi.org/10.14195/2184-9781_6_0

do juiz, descurando o complexo de atividades que este efetivamente realiza. Segundo os autores, o horizonte judiciário tem-se diversificado e ganho um universo decisional associado à ideia do justo e do equitativo - um juiz guiado por diversas fontes de regulação (regras legais, princípios, consciência), por diferentes objetos (jurisprudencial, processual, de apuramento da verdade, sancionatório, das liberdades, tutelar) e por distintas intensidades de intervenção.[32] Esta diversificação é anterior à revolução digital e não se explica por ela; mas é precisamente sobre este pano de fundo, já de si complexo, que a inteligência artificial vem incidir.

No plano constitucional português, Gomes Canotilho situa a posição do poder jurisdicional de forma distinta da dos demais órgãos de soberania: não pautada por uma representatividade igual à destes, mas fundada e legitimada na administração da justiça em nome do povo - nexo que a Constituição exprime através da noção de delegação (artigo 202.º) - e no acesso imediato e direto à satisfação de princípios constitucionais, como a apreciação da constitucionalidade (artigo 204.º).[33] Antoine Garapon, ainda antes, ao refletir sobre justiça e democracia, convidava a reconhecer que o real papel do juiz não é tomar o lugar do político, mas resistir ao risco de implosão da democracia, permanecendo guardião das promessas inscritas nas leis da república.[34] Mais recentemente, os coordenadores do European Yearbook of Constitutional Law de 2019 sublinhavam que a independência judicial não se pode confundir com isolamento, antes se realizando como uma relação institucional que legitima o poder dos tribunais perante os cidadãos e as demais instituições do Estado.[35]

Do que resulta este breve percurso? Que a atividade jurisdicional ultrapassa em muito a vertente cognitiva ou a aplicação mecânica e reprodutiva da lei - o modelo, hoje ultrapassado, do juiz como boca da lei. Ela é antes a realização jurisdicional do direito: um processo em que o juiz, perante um caso concreto, interpreta factos e normas num espaço de ponderação e de decisão criativa, fundamentado numa estrutura lógica, linguística, cognoscitiva e argumentativa, e que se desdobra em três dimensões que se implicam mutuamente - a

32Garapon, A., Perdriolle, S., Bernabé, B., & Kadri, C. (2013). La prudence et l'autorité: L'office du juge au XXIe siècle [Rapport]. Paris: Institut des Hautes Études sur la Justice (IHEJ), pp. 16-23.

33Canotilho, J. J. G. (2003). Direito constitucional e teoria da constituição (7.ª ed.). Coimbra: Almedina, pp. 657-658.

34Garapon, A. (1996). Le gardien des promesses: Justice et démocratie. Paris: Odile Jacob, pp. 143-156.

35Hirsch Ballin, E., van der Schyff, G., & Stremler, M. (Eds.). (2020). European Yearbook of Constitutional Law 2019: The Role of the Judiciary in Contemporary Constitutional Systems. Haia: T.M.C. Asser Press, pp. 3-12.

intenção de resolução de controvérsias, a prática judicativa e a organização ou estrutura que a suporta[36]. É contra este pano de fundo, e não contra uma imagem simplificada do juiz-autómato, que a relação entre a inteligência artificial e a atividade jurisdicional deve ser pensada.

IA: uma tecnologia de predição, não de julgamento

A inteligência artificial, no estádio atual do seu desenvolvimento, é uma tecnologia de predição - não de julgamento. Processa volumes de dados inacessíveis ao juiz, identifica padrões jurisprudenciais e pode reduzir vieses cognitivos gerados pela sobrecarga de informação. Não aplica, todavia, cláusulas gerais; não distingue *ratio decidendi* dos *obiter dicta*; não avalia a credibilidade de uma testemunha; não fundamenta com intencionalidade e consciência moral. Eljas e Tuula Linna, num estudo recente sobre os desafios do raciocínio jurídico gerativo, confirmam esta fronteira ao decompor os requisitos nucleares do raciocínio judicativo - a seleção do quadro legal aplicável entre jurisdições, a construção de argumentos assentes na doutrina das fontes do direito, a distinção entre *ratio decidendi* e *obiter dicta*, a resolução da ambiguidade própria de cláusulas gerais como a razoabilidade - e ao demonstrar que os mecanismos de otimização hoje disponíveis, incluindo a geração aumentada por recuperação (RAG) e os sistemas multiagente, colmatam desafios técnicos específicos, mas não os mais relevantes, que continuam a exigir discricionariedade e uma fundamentação transparente e justificável.[37]

Arthur Dyevre e Ahmad Shahvaroughi, numa revisão sintética que cruza a psicologia, a economia, o direito e a criminologia, chegam a uma conclusão convergente a partir de outro ângulo: a evidência empírica disponível sobre o impacto das ferramentas de apoio à decisão por inteligência artificial nas decisões de pré-julgamento e de sentenciamento é, quando muito, modesta, ao mesmo tempo que revela lacunas relevantes sobre o modo como os juízes navegam ambientes decisórios ruidosos e como as características individuais influenciam a resposta ao conselho algorítmico.[38] O retrato empírico que emerge destes dois estudos é

36 Assim, Coelho Nuno (2024). Ob. cit., em diversos passos do texto do livro.
37Linna, E., & Linna, T. (2026). Challenges for generative AI in legal reasoning. Discover Artificial Intelligence, 6, artigo 170.
38Dyevre, A., & Shahvaroughi, A. (2026). Man and machine: Artificial intelligence and judicial decision making (preprint). Lovaina: Leuven Centre for Empirical Jurisprudence.

convergente com o quadro epistemológico traçado no ponto anterior: a máquina humeana prevê; não conclui.

O risco do deskilling e o «conformismo judiciário»

Esta fronteira, precisamente porque não é sempre visível no imediato, gera um risco de sentido inverso. Anabela Miranda Rodrigues designou por conformismo judiciário aquilo que a literatura anglo-saxónica tem vindo a chamar *deskilling*: a tendência dos operadores humanos para aceitarem acriticamente as conclusões dos mecanismos preditivos, abdicando da responsabilidade hermenêutica em favor de uma celeridade processual muitas vezes ilusória.[39] O risco não é abstrato, nem exclusivo da magistratura portuguesa: é a mesma rendição cognitiva e a mesma epistemia já descritas no ponto anterior, agora observadas no local exato onde os seus efeitos mais importam - o momento em que se decide da liberdade, do património ou dos direitos de uma pessoa concreta.

Dois contributos de matriz mais antropológica do que jurídica ajudam a explicar por que motivo esta cedência é tão insidiosa. Thomas Fuchs sustenta que a inteligência da decisão em ambiente social se partilha e se fundamenta através de narrativas empáticas imersas em complexidade humana aguda, sem transitividade possível para meras extrapolações estatísticas lineares.[40] Neil Lawrence, por seu turno, situa esta mesma irredutibilidade na diferença entre a informação que um sistema processa e a experiência corporizada e situada de que um ser humano dispõe para lhe atribuir sentido.[41] Delegar silenciosamente o ato jurisdicional na máquina não é, à luz destes contributos, uma mera transferência de tarefas: é o recuo de uma responsabilidade pessoal e institucional que não tem equivalente algorítmico.

Do «juiz robot» ao «juiz cyborg»: o modelo incremental

É este o pano de fundo em que se situa a distinção, já desenvolvida anteriormente, entre um modelo substitutivo e um modelo incremental de integração tecnológica na atividade

39Rodrigues, A. M. (Coord.). (2020). Inteligência artificial no direito penal. Coimbra: Almedina.
40Fuchs, T. (2021). In defense of the human being: Foundational questions of an embodied anthropology. Oxford: Oxford University Press.
41Lawrence, N. D. (2024). The atomic human: Understanding ourselves in the age of AI. Londres: Allen Lane.

jurisdicional.[42] No primeiro, a inteligência artificial toma o lugar do juiz - a imagem, tornada célebre por Eugene Volokh, de um *Chief Justice Robot* cujas decisões, sendo persuasivas e consistentes, mereceriam aceitação independentemente da sua origem não humana; uma tese que, mesmo quando formulada com todo o rigor argumentativo, este texto rejeita pelas razões já expostas.[43] No segundo modelo - o modelo incremental - a tecnologia amplifica as capacidades do julgador sem substituir o seu discernimento: a imagem, aqui preferida, não é a do juiz *robot*, mas a do juiz *cyborg*, munido e capacitado pela técnica sem lhe ser reduzido. O princípio estruturante deste modelo incremental pode formular-se como *Human IN Command*: o humano não como variável de supervisão residual, acionada apenas quando a máquina falha, mas como princípio constitutivo da própria decisão, presente do início ao fim do processo judicativo. É esta a arquitetura de responsabilidade que o quarto ponto deste texto procurará agora decompor, distinguindo os diferentes níveis de intervenção tecnológica segundo a sua proximidade crescente com o ato de julgar.

A atividade a montante: escritórios e bastidores dos tribunais

A intensidade, o alcance e os limites de utilização da inteligência artificial na atividade jurisdicional não se determinam apenas no momento processual em sentido estrito. Determinam-se, em larga medida, a montante - na atividade desenvolvida nos escritórios de advocacia e nos bastidores organizativos dos tribunais, onde se preparam os articulados, se gere a informação processual e se estrutura o fluxo de trabalho que antecede qualquer intervenção jurisdicional. Nos escritórios de advocacia, o recurso a ferramentas de inteligência artificial generativa na elaboração de peças processuais, na pesquisa jurisprudencial e na revisão documental tem um efeito de propagação inevitável: o que entra no processo já vem filtrado, estruturado e, por vezes, gerado por sistemas cuja metodologia, fiabilidade e exposição a riscos de alucinação o magistrado desconhece.

Na organização e gestão dos tribunais, a adoção de sistemas de inteligência artificial para triagem, distribuição, classificação documental e gestão de prazos condiciona estruturalmente o que chega ao gabinete do juiz e em que formato, antes mesmo de qualquer decisão jurisdicional ser tomada. A questão da admissibilidade e do controlo da inteligência

42Coelho, N. (2024). Ob. cit., caps. 5 e 12, pp. 664-702.
43Volokh, E. (2019). Chief Justice robots. Duke Law Journal, 68(6), 1135-1192.

artificial no processo não pode, por isso, confinar-se à atuação do juiz: exige uma reflexão sobre deveres de transparência e de verificação que recaem também sobre os mandatários e sobre os próprios serviços de gestão processual - reflexão que o quarto ponto deste texto começará precisamente por este nível, o mais distante do ato de julgar e, por isso mesmo, o mais frequentemente subestimado.

Uma reconfiguração constitucional, não uma atualização técnica

The Cambridge handbook of AI and Technologies in Courts, publicado neste ano de 2026 sob a coordenação de Monika Zalnieriute e Agne Limante, sintetiza num só argumento o que este ponto procurou demonstrar por vias diversas: a introdução da inteligência artificial nos tribunais não é uma mera atualização tecnológica, mas uma reconfiguração constitucional e um novo apelo à responsabilização judicial - e a maior parte da discussão sobre inteligência artificial e tribunais tem-se concentrado, de forma injustificadamente estreita, nas ferramentas de decisão, quando os tribunais desempenham um conjunto de funções societais muito mais amplo do que a mera resolução de litígios.[44] É precisamente esta amplitude - da gestão processual à decisão final, do escritório de advocacia ao gabinete do juiz - que o quarto ponto deste texto procurará agora cartografar, distinguindo os níveis de intervenção tecnológica segundo o seu grau de proximidade com o núcleo irredutível da função de julgar.

IV. OS NÍVEIS DE INTERVENÇÃO TECNOLÓGICA NA ATIVIDADE JURISDICIONAL

O ponto anterior distinguiu o modelo substitutivo do modelo incremental e anunciou uma cartografia dos níveis de intervenção tecnológica, organizada segundo o seu grau de proximidade com o ato de julgar. É essa cartografia que este quarto ponto desenvolve - não como um exercício classificatório gratuito, mas porque cada nível convoca ferramentas,

44Zalnieriute, M., & Limante, A. (Eds.). (2026). The Cambridge handbook of AI and technologies in courts. Cambridge: Cambridge University Press.

princípios de supervisão e riscos específicos, e só a sua distinção permite formular, mais adiante, respostas regulatórias e formativas proporcionadas a cada um.[45]

A utilidade da inteligência artificial na atividade judicial organiza-se, deste modo, em três níveis de proximidade crescente com o ato de julgar - cada um com as suas ferramentas típicas, o seu princípio de supervisão e o seu risco específico:

	Nível I - Otimização da Gestão	**Nível II - Apoio à Decisão**	**Nível III - Reflexão Crítica**
	Tribunais e escritórios	**Apoio qualificado ao juiz**	**Aprimoramento da decisão**
Ferramentas	Cronogramas e *data mining*; triagem, distribuição, bots; anonimização e tradução automática; classificação de processos e peças	Pesquisa jurisprudencial (*argument mining*); RAG sobre fontes verificadas; avaliação preditiva do mérito; sugestão de sumários	Destilação de pesquisa; desenviesamento cognitivo; análise de cenários; geração de opções
Supervisão	*Human ON-the-loop*	*Human IN-the-loop*	*Human IN Command*
Riscos específicos	*Shadow AI*	Alucinação	*Prompt injection* *Deskilling* *Grounding*

Nível I - Otimização da gestão: a montante, mas não neutro

O primeiro nível retoma o que o ponto anterior designou por atividade a montante: cronogramas, triagem e distribuição de processos, anonimização, tradução automática, classificação documental. É tentador tratar este nível como puramente técnico e desprovido de relevância jurisdicional - afinal, nenhuma destas ferramentas decide, sugere

45 Os três níveis não são estanques: a fronteira entre eles desloca-se com a maturidade tecnológica e institucional de cada tribunal, e uma mesma ferramenta pode migrar de nível consoante o modo como é configurada e supervisionada.

fundamentação ou substitui juízo algum. Jürgen Habermas, cuja obra sobre ciência e técnica como ideologia continua a merecer leitura atenta, alertava precisamente para o risco inverso: em sociedades industriais avançadas, a ciência e a tecnologia tendem a converter questões políticas em questões técnicas, retirando-as do espaço da deliberação pública sob o argumento de que são demasiado complexas para o cidadão comum.[46] Aplicado à organização judiciária, o aviso é direto: a escolha de um sistema de triagem processual, de um algoritmo de distribuição de processos ou de um critério de classificação documental não é uma decisão neutra de eficiência informática - é uma decisão institucional com efeitos sobre quem acede a quê, quando e em que ordem, e que, precisamente por se apresentar como técnica, tende a escapar ao escrutínio que mereceria.

É por isso que o princípio de supervisão adequado a este nível - *Human ON-the-loop*, o humano sobre o circuito, capaz de o interromper ou corrigir, mas não necessariamente presente em cada operação - não pode ser confundido com ausência de responsabilidade institucional. O risco específico deste nível, a *shadow AI*, ilustra bem o problema: quando funcionários judiciais ou mandatários recorrem a ferramentas de inteligência artificial generativa não homologadas, à margem dos canais oficiais, para tarefas de gestão ou de redação, a instituição perde precisamente aquela capacidade de supervisão que o modelo pressupõe. A montante do ato de julgar, a governação tecnológica não é um luxo administrativo: é a condição de possibilidade de todos os níveis que se seguem.

Nível II - Apoio à decisão: o que a evidência empírica mostra

É neste segundo nível que se concentra o essencial da investigação empírica sobre inteligência artificial e decisão judicial, sobretudo no domínio da avaliação de risco em processo penal - cauções, sentenciamento, liberdade condicional. Jon Kleinberg e coautores, num estudo de referência sobre centenas de milhares de decisões de fiança, treinaram um modelo de aprendizagem automática para prever o risco de incumprimento e simularam os seus efeitos: a substituição das decisões humanas pela recomendação do modelo permitiria, nas suas estimativas, reduções de até 24,7% na taxa de criminalidade e de até 41,9% na taxa de encarceramento, sem agravar - e nalguns casos reduzindo - as disparidades raciais

46Habermas, J. (1970). Toward a rational society: Student protest, science, and politics (J. J. Shapiro, Trad.). Boston: Beacon Press. (Ensaio original, "Technik und Wissenschaft als 'Ideologie'", 1968.)

observadas.[47] Trata-se de um resultado assinalável, mas isolado: quando Ben-Michael, Kosuke Imai e coautores conduziram o primeiro ensaio aleatorizado controlado que testou o impacto real de um sistema de avaliação de risco (a *Public Safety Assessment*) sobre decisões de juízes em exercício - com uma amostra superior a 1.800 casos - o resultado foi nulo: as taxas de falsos positivos e de falsos negativos para a reincidência não diferiram de modo estatisticamente significativo entre o grupo com acesso ao sistema e o grupo de controlo.[48]

Esta divergência entre a simulação e o ensaio controlado é instrutiva por si só: sugere que o desempenho teórico de um modelo preditivo - medido, tipicamente, por valores de área sob a curva (AUC) que, numa revisão sistemática de seis instrumentos de avaliação de risco pré-julgamento, se situam entre 0,64 e 0,73 - está longe de se traduzir automaticamente em ganhos reais quando o sistema é apenas um entre muitos *inputs* que o juiz pondera, e não o substituto da sua decisão. A literatura sobre o viés de automação (*automation bias*) ajuda a compreender esta discrepância: a disponibilidade de uma recomendação algorítmica altera o comportamento humano de formas que nem sempre são as esperadas, nem sempre na direção da sobreconfiança acrítica, nem sempre na direção da rejeição sistemática.[49]

No plano português, esta cautela ganha densidade específica com os alertas de Nádia Alberto sobre as lacunas do quadro legislativo nacional perante o recurso a *big data* e a algoritmos preditivos no processo penal, e com o trabalho coordenado por Manuel Monteiro Guedes Valente sobre a admissibilidade e a valoração da prova obtida por inteligência artificial, o reconhecimento facial e as novas fronteiras da criminalidade digital - trabalhos que convergem na conclusão de que a interpretação da lei criminal perante o caso concreto exige uma valoração ética e humana incompatível com o determinismo de máquina.[50] É este o nível em que o princípio *Human IN-the-loop* - o humano dentro do circuito, presente em cada

47Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. The Quarterly Journal of Economics, 133(1), 237-293.

48Ben-Michael, E., Greiner, D. J., Huang, M., Imai, K., Jiang, Z., & Shin, S. (2025). Does AI help humans make better decisions? A statistical evaluation framework for experimental and observational studies. Proceedings of the National Academy of Sciences, 122(38), e2505106122.

49Desmarais, S. L., Zottola, S. A., Duhart Clarke, S. E., & Lowder, E. M. (2021). Predictive validity of pretrial risk assessments: A systematic review of the literature. Criminal Justice and Behavior, 48(4), 398-420; Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. Journal of the American Medical Informatics Association, 19(1), 121-127.

50Alberto, N. (2024). Algoritmos e big data a partir do sistema de justiça criminal português: Contributos para uma justiça automatizada. Coimbra: Almedina; Valente, M. M. G. (Coord.). (2024). Ciências Criminais e Inteligência Artificial.

decisão relevante - encontra a sua justificação mais forte, e em que o risco de alucinação, fáctica ou jurisprudencial, mais diretamente compromete a fundamentação da decisão final.

Nível III - Reflexão crítica: o Método Dialógico Multirresposta

O terceiro nível não pretende que a inteligência artificial decida, nem sequer que proponha uma única solução: pretende que amplifique a reflexão crítica do julgador. Neste nível pretende-se o aprimoramento da reflexão crítica, com a destilação de pesquisa, o desenviesamento cognitivo e análise dos cenários possíveis em face das consequências de cada uma das decisões possíveis no caso. A inteligência artificial pode, aqui, amplificar o raciocínio, mas a consciência moral, a intencionalidade e a responsabilidade decisional devem permanecer irredutivelmente no juiz (*Human IN Comand*). Os riscos envolvidos e que têm vindo a ser catalogados como podem ser prevenidos e afastadas as suas negatividades?

O Método Dialógico Multirresposta, desenvolvido pelo desembargador Marco Villas Boas no âmbito do Projeto Hórus - uma cooperação entre o Grupo de Pesquisa Smart Cities BR da Escola de Artes, Ciências e Humanidades da Universidade de São Paulo, o Colégio de Presidentes de Escolas Estaduais da Magistratura e a Escola Superior da Magistratura Tocantinense - oferece um modelo operacional para este nível, estruturado em quatro eixos: a interação dialógica qualificada, que decompõe o raciocínio jurídico em premissas menores através do refinamento iterativo de comandos; a exploração plural, que obriga o sistema a gerar múltiplos caminhos argumentativos em vez de uma resposta única, precisamente para contrariar o enviesamento de confirmação do operador humano; a filtragem crítica transparente, em que o magistrado audita as fontes normativas e jurisprudenciais invocadas; e a preservação da autonomia decisória, que mantém a primazia do julgamento vinculada de forma absoluta à prudência do agente humano.[51]

Este modelo não surge isolado na literatura internacional. Já em 1998, numa coletânea de referência sobre aplicações judiciais da inteligência artificial, Giovanni Sartor e L. Karl Branting sublinhavam que a investigação nesta área sempre recusou o sonho - ou pesadelo - de um juiz automático, orientando-se antes para instrumentos de apoio que preservassem a discricionariedade decisória; e, na mesma coletânea, Michele Taruffo alertava para a

51Villas Boas, M. (2024). O uso ético e estratégico da inteligência artificial generativa no judiciário: O método dialógico multirresposta. São Paulo: EACH/USP; COPEDEM; ESMAT.

impossibilidade conceptual de reduzir o raciocínio prudencial e valorativo do julgador a modelos formais de computação lógica.[52] O que separa o nível III dos dois anteriores não é, por isso, a sofisticação técnica da ferramenta, mas a sua função: já não gerir nem prever, mas confrontar o julgador com alternativas que ele próprio deve escrutinar. É aqui que o princípio *Human IN Command* adquire o seu sentido pleno - o humano não como validador final de um produto acabado, mas como autor de um processo que a máquina apenas instrui - e é também aqui que o risco de *prompt injection*, a manipulação da interação através de instruções ocultas ou maliciosas inseridas no próprio material processado, mais gravemente ameaça a integridade do processo dialógico que o método pressupõe.

O «humano no circuito» não garante, por si, segurança

Os três princípios de supervisão que atravessam esta tipologia - *Human ON-the-loop, IN-the-loop, IN Command* - partilham uma premissa que importa agora questionar: a de que a presença formal de um humano no circuito basta para neutralizar os riscos da automação. Gali Noti e coautores, num estudo sobre sistemas de apoio à decisão em que o humano aprende através de interações repetidas com o algoritmo, mostram que esta premissa pode falhar de um modo subtil.[53] Quando o algoritmo é desenhado para maximizar a precisão da decisão segundo o seu próprio modelo, e não para desenvolver a competência do operador humano, gera-se uma tensão estrutural: fornecer ao humano apenas a informação mais precisa no imediato pode reduzir, a prazo, a sua capacidade de compreender autonomamente o domínio em que opera - precisamente o quadro em que, se o sistema disponibilizar uma saída que o utilizador não consegue verificar ou compreender por não ter acompanhado o desenvolvimento do seu próprio conhecimento de domínio, o humano permanece formalmente no circuito sem nele funcionar como salvaguarda substantiva.

Transposta para a atividade jurisdicional, esta advertência atinge todos os três níveis, mas com intensidade crescente do primeiro para o terceiro: um juiz que aprendeu a confiar cegamente nas sugestões de pesquisa jurisprudencial do nível II, sem desenvolver

52Sartor, G., & Branting, L. K. (Eds.). (1998). Judicial applications of artificial intelligence. Dordrecht: Kluwer Academic Publishers; Taruffo, M. (1998). Judicial decisions and artificial intelligence. Artificial Intelligence and Law, 6(2-4), 311-324.

53Noti, G., Donahue, K., Kleinberg, J., & Oren, S. (2025). AI-assisted decision making with human learning. Proceedings of the 26th ACM Conference on Economics and Computation (EC '25), 38-39.

paralelamente a competência para as questionar, pode formalmente estar *IN-the-loop* e substantivamente já não o estar. A distinção entre os três níveis de intervenção tecnológica, por isso, não substitui a exigência formativa que o sexto ponto deste texto desenvolverá: é a sua precondição. Sem uma cultura epistemológica tecnológica que mantenha viva a competência hermenêutica em cada nível, a arquitetura de supervisão aqui descrita corre o risco de se tornar uma ficção institucional tranquilizadora.

Uma tipologia ancorada em soft law, não em abstração doutrinal

Esta tipologia de três níveis não é uma construção puramente doutrinal: converge com o quadro regulatório que, em Portugal e na Europa, já vincula a atividade judicial. As Recomendações para o Uso de Inteligência Artificial na Atividade Judicial, aprovadas pelo Conselho Superior da Magistratura em abril de 2026, consagram o carácter estritamente auxiliar destas ferramentas, o controlo humano efetivo sobre a decisão, a responsabilização plena do magistrado e o dever de menção do seu uso - princípios que correspondem, ponto por ponto, aos três níveis de supervisão aqui distinguidos.[54] No plano europeu, a Carta Ética Europeia sobre o Uso da Inteligência Artificial nos Sistemas Judiciais e no seu Ambiente, adotada pela Comissão Europeia para a Eficiência da Justiça, assenta no mesmo princípio de centralidade humana, impondo a retenção de controlo pleno pelo utilizador em todos os níveis de intervenção.[55] É com esta arquitetura de níveis, princípios de supervisão e riscos específicos estabelecida que o quinto ponto deste texto pode agora avançar para a dimensão propriamente hermenêutica da decisão judicial - o modelo IRAC e os limites que a tecnologia, em qualquer dos três níveis descritos, não ultrapassa.

V. A DIMENSÃO HERMENÊUTICA. O MODELO IRAC E OS LIMITES DA TECNOLOGIA

O segundo ponto deste texto introduziu, a propósito de Kant, a distinção entre o juízo determinante e o juízo reflexionante, e sustentou que a decisão judicial se aproxima

54Conselho Superior da Magistratura de Portugal. (2026). Recomendações para o Uso de Inteligência Artificial na Atividade Judicial [GAIA 2025/GAVPM/5469]. Lisboa: CSM.

55CEPEJ - Comissão Europeia para a Eficiência da Justiça. (2018). Carta Ética Europeia sobre o Uso da Inteligência Artificial nos Sistemas Judiciais e no seu Ambiente. Estrasburgo: Conselho da Europa.

estruturalmente deste último. O terceiro e o quarto pontos mostraram como essa distinção se projeta na atividade concreta do juiz e nos diferentes níveis de intervenção tecnológica. Chega agora o momento de descer ao núcleo mais denso e identitário dessa atividade - a dimensão hermenêutica da decisão judicial - e de o fazer com o rigor analítico que o tema exige, através do designado modelo IRAC (*Issue, Rule, Application, Conclusion*) e da pergunta que a este texto mais interessa: onde, exatamente, a tecnologia encontra o seu limite, e porquê.

Hermenêutica filosófica: a compreensão como fusão de horizontes

A tradição hermenêutica que se inicia com Heidegger e Dilthey e amadurece com Gadamer contesta um pressuposto que a modernidade tomou por adquirido: o de que existe uma separação nítida entre o sujeito cognoscente e o objeto conhecido, pela qual aquele captaria a informação proveniente deste sem a contaminar. Sempre que se aborda cognitivamente um objeto, sustenta esta tradição, já o fazemos dentro de um horizonte de compreensão definido pela tradição, pela história e pela cultura - e, no caso do jurista, também pela formação profissional e pela experiência acumulada de casos anteriores. Arthur Kaufmann formulou esta tese de modo particularmente incisivo: a hermenêutica opõe-se ao conceito objetivista de conhecimento, superando o esquema sujeito/objeto aplicado à compreensão, porque o intérprete não se limita a representar passivamente o objeto na sua consciência, mas desempenha um papel conformador ativo naquilo a que se chama aplicação do direito.[56] [57]

Karl Larenz precisou esta ideia a propósito do jurista em particular: a sua pré-compreensão resulta de um longo processo de aprendizagem, que integra tanto os conhecimentos adquiridos na formação como as experiências profissionais e extraprofissionais posteriores, sobretudo as que respeitam a factos e a contextos sociais. É esta mesma tese que, como já se viu, Baptista Machado defendeu a propósito do círculo hermenêutico entre facto e norma, e que Castanheira Neves radicalizou ao fazer do caso concreto, e não da norma abstrata, o *prius*

56Assim, em Silva, R. M. M. T. da (2025). Interpretação artificial: Sobre a possibilidade de os novos sistemas de inteligência artificial fazerem interpretação jurídica. Coimbra: Almedina, pp. 103-105.

57Kaufmann, A. (2002). Filosofia do direito, teoria do direito, dogmática jurídica. Em A. Kaufmann & W. Hassemer (Orgs.), Introdução à filosofia do direito e à teoria do direito contemporâneas (M. Keel & M. Seca de Oliveira, Trad.). Lisboa: Fundação Calouste Gulbenkian.

metodológico do pensamento jurídico. Compreender um caso jurídico não é, nesta ótica, aplicar mecanicamente uma grelha a um input neutro: é um processo em que o próprio horizonte do intérprete - a sua pré-compreensão, na terminologia de Larenz e de Gadamer - conforma ativamente o sentido do que é compreendido.

Uma objeção previsível merece ser reconhecida com cuidado, e não liminarmente afastada: se toda a compreensão parte de pré-juízos, não estará a hermenêutica filosófica a legitimar o preconceito? A objeção tem mérito enquanto alerta - a pré-compreensão de um julgador pode, de facto, incorporar enviesamentos que a Constituição proíbe, como o preconceito quanto à origem social ou aos antecedentes de um arguido. Mas esta possibilidade não invalida a tese hermenêutica; exige, antes, que a pré-compreensão seja permanentemente sujeita a controlo racional e a fundamentação, e não que se finja a sua inexistência. É precisamente este ponto de equilíbrio - pré-compreensão inevitável, mas sujeita a escrutínio racional - que separa a hermenêutica filosófica de um subjetivismo puro, e que, como se verá, nenhum sistema de inteligência artificial está em condições de replicar, porque não possui pré-compreensão alguma a escrutinar.

Uma compreensão multifacetada: a hermenêutica como vertente, como círculo e como reserva do humano

A hermenêutica filosófica que se acabou de expor não esgota, todavia, o modo como a reflexão jurídica portuguesa contemporânea trata o problema hermenêutico - e a globalidade da questão exige que se reconheça, antes de mais, esta pluralidade de registos. Já na apresentação da obra que serve de referência sistemática a este texto, a hermenêutica surge não como objeto de análise, mas como descrição do próprio método de construção do livro: a «lógica circular e a textura sedimentar do trabalho», nas palavras deste mesmo autor, remetem para uma abordagem hermenêutica que explica o porquê de sucessivas aproximações às distinções nele assumidas.[58] A circularidade, longe de ser uma fragilidade metodológica, é aqui assumida como inevitável - e antecipa, em registo autobiográfico, a estrutura que a hermenêutica voltará a assumir noutros planos do mesmo percurso teórico.

Num segundo registo, a hermenêutica surge não como teoria envolvente da interpretação jurídica, mas como uma entre várias modalidades de racionalidade jurídica - a par das

58Coelho, Nuno (2024). Ob. cit., p. 29.

vertentes procedimental, tópico-retórica, argumentativa, teleológica e crítica, na enumeração que Castanheira Neves propôs e que se mantém estável ao longo de mais de quinhentas páginas de desenvolvimento teórico e de análise da decisão judicial concreta.[59] José Lamego precisa o conteúdo específico desta vertente, definindo-a como um saber prático (*knowing how*) centrado na aplicação do direito e na compreensibilidade a partir do ponto de vista interno do intérprete - uma formulação que aproxima a hermenêutica, mais do que de uma doutrina geral da interpretação de textos, de uma competência prática do julgador.[60] Esta precisão importa para o argumento deste texto: nenhum sistema de inteligência artificial possui um «ponto de vista interno», precisamente porque não é um sujeito prático, mas um mecanismo de correlação estatística sobre símbolos.

Num terceiro registo, a circularidade hermenêutica desloca-se do plano da compreensão individual para o plano sistémico e para o plano decisório. No primeiro, fala-se de uma ponderação das consequências das várias propostas de solução por via da *circularidade hermenêutica*, enquanto mecanismo de auto-ajustamento do próprio sistema jurídico enquanto estrutura autopoiética - retomando, num registo mais técnico, a tese sobre a autopoiese do direito já desenvolvida no segundo ponto deste texto.[61] No segundo, fala-se de um círculo hermenêutico entre intérprete e norma, que Karl Larenz situou como «um modo de pensamento de grande alcance na jurisprudência» e que Michele Taruffo reconduziu à operação, tantas vezes complexa, pela qual o juiz identifica a premissa de direito da decisão mediante opções hermenêuticas orientadas por cânones interpretativos - sem prejuízo de, na mesma operação, realizar juízos de valor assentes em opções éticas, políticas e económicas.[62] A hermenêutica não é, pois, um estádio que se ultrapassa uma vez identificada a norma aplicável: acompanha, em circularidade, todo o processo decisório.

É, por fim, num quarto e último registo - o mais diretamente relevante para o tema deste texto - que a hermenêutica surge explicitamente convocada perante a inteligência artificial. Ao delimitar a reserva da integridade do juízo humano face à automação, a obra que aqui se segue enumera as componentes desse juízo «*em toda a sua amplitude argumentativa,*

59Castanheira Neves, A. (1993). Ob. cit., pp. 70-81; Coelho, Nuno (2024). Ob. cit., pp. 152-153, 726-727.
60Lamego, J. (1990). Hermenêutica e jurisprudência: Análise de uma receção. Lisboa: Fragmentos, pp. 157-165.
61Coelho, Nuno (2024). Ob. cit., p. 126.
62Larenz, K. (1989). Ob. cit., pp. 242-243; Taruffo, M. (1998). Ob. cit., pp. 311-324; Coelho, Nuno (2024). Ob. cit., pp. 727-728.

dedutiva, inferencial, interpretativa, hermenêutica, argumentativa, analógica, que convoca princípios e pondera valores», contrapondo-lhe, do lado da máquina, apenas «a apreciação matemática e lógica, estatística, padronizada, se bem que sofisticada, racional e informada, mas sempre artificial».[63] A hermenêutica figura aqui, lado a lado com a argumentação, a dedução, a inferência e a analogia, entre os traços do juízo humano que este texto sustenta, desde a sua introdução, não poderem ser delegados sem perda de identidade na própria função de julgar.

Método declarado, vertente entre vertentes, padrão circular replicado em vários planos, traço da reserva do humano - a hermenêutica que este ponto do texto convoca não é, por isso, uma doutrina única e fechada, mas um problema que atravessa transversalmente a realização do direito. É contra este pano de fundo multifacetado, e não contra uma noção estreita de «interpretação de textos», que importa agora situar a arquitetura mais específica da abertura, da moldura e do fechamento - e, através dela, o próprio modelo IRAC.

A jurisdição entre a abertura, a moldura e o fechamento

A hermenêutica filosófica não esgota, todavia, o modo como a metodologia jurídica portuguesa contemporânea tem pensado o alcance da tarefa realizadora do direito. Recorde-se, do segundo ponto deste texto, a definição de jurisdição como a atividade exercida pelos juízes destinada à revelação, extrinsecação e aplicação do direito num caso concreto. O alcance dessa tarefa oscila entre dois polos que a tradição designa por fechamento e abertura: a via interpretativa mais fechada, ou legalista, tende a salvaguardar o monopólio legislativo da criação do direito; a via mais aberta, ou realista, tende a deslocar esse monopólio para o poder judicial, ou mesmo para aquilo a que se tem chamado a comunidade dos juristas.[64]

Entre estes dois extremos situa-se a metáfora da moldura, que Hans Kelsen introduziu para descrever a pluralidade de sentidos possíveis que a tarefa interpretativa admite dentro dos limites fixados pela norma.[65] Ronald Dworkin retomou a mesma imagem com um sentido distinto: a moldura, para o autor norte-americano, é o modelo da relação entre os princípios jurídicos e as decisões judiciais, em que aqueles desempenham um papel delimitador da ação

63Coelho, Nuno (2024). Ob. cit., p. 699.
64Coelho, Nuno (2024). Ob. cit., pp. 130-131, 165.
65Kelsen, H. (2005). Teoria pura do direito (J. B. Machado, Trad., 7.ª ed. alemã), pp. 350-352. Coimbra: Almedina.

do intérprete, conferindo coerência e significado a estas últimas e emprestando ao sistema a integridade que Dworkin lhe reivindicava.[66] José Manuel Aroso Linhares propôs decompor esta problemática em três núcleos distintos, mas interligados: a *jurisdictio* como organização ou estrutura, a *jurisdictio* como intenção de resolução de controvérsias e a *jurisdictio* como dimensão da *praxis*.[67]

A moldura, nesta leitura mais enriquecida, não é apenas linguística - a pluralidade de sentidos que um texto normativo comporta - mas institucional: é a jurisdição enquanto facto institucional, segundo os padrões que a própria comunidade jurídica considera válidos.[68] Esta dupla natureza da moldura - hermenêutica e institucional - é o que verdadeiramente importa reter para a questão que este texto persegue: um sistema de inteligência artificial que produza um resultado interpretativo plausível fá-lo, na melhor das hipóteses, dentro dos limites linguísticos da moldura kelseniana; não tem, porém, acesso à moldura institucional que Dworkin e Aroso Linhares descrevem, porque essa moldura depende de uma validação comunitária e axiológica que nenhum sistema técnico pode, por si, produzir. É esta mesma distinção, ainda que noutros termos, que Ricardo Tavares da Silva explora ao separar o processo interpretativo do resultado que dele emerge: um sistema de inteligência artificial pode aproximar-se do resultado que a moldura comporta, sem que essa aproximação signifique acesso ao processo institucional que lhe confere validade.[69]

Do esquema clássico aos modelos jurisdicionais contemporâneos: uma equação com o modelo IRAC

Antes de o modelo IRAC ser importado para o debate português sobre inteligência artificial e decisão judicial, a metodologia jurídica nacional já dispunha da sua própria decomposição da tarefa realizadora do direito - e é essa decomposição, e não apenas a arquitetura kelseniano-dworkiniana já exposta, que permite equacionar rigorosamente o modelo IRAC com as metodologias jurisdicionais que o precedem. O esquema clássico de realização do direito assentava na compartimentação do processo interpretativo-aplicativo e

66Dworkin, R. (1978). Taking rights seriously. Cambridge, MA: Harvard University Press, pp. 28-80.
67Linhares, J. M. A. (2010). Jurisdição, diferendo e «área aberta»: A caminho de uma teoria do direito como moldura. Em *Ars Ivdicandi - Estudos em Homenagem ao Prof. Doutor Jorge de Figueiredo Dias*, Vol. IV, pp. 450-453. Coimbra: Coimbra Editora.
68Coelho, N. (2024). Ob. cit., p. 192.
69Silva, R. M. M. T. da (2025). Ob. cit., pp. 139-140.

no método de subsunção: a determinação da fonte relevante, a sua interpretação, a integração de eventuais lacunas - «*admitidas com dificuldade e em última ratio*» - a delimitação da matéria de facto, a sua qualificação jurídica e, por fim, a aplicação; como elementos auxiliares desta tarefa, distinguiam-se os elementos gramaticais ou literais, os lógicos e sistemáticos, os históricos e os teleológicos, hoje enquadrados, embora de modo tendencialmente unitário, nos artigos 9.º e 10.º do Código Civil.[70]

Este esquema foi superado, décadas antes de qualquer discussão sobre inteligência artificial, por um modelo em *continuum*, no qual interpretação, integração e ponderação das consequências passaram a ser pensadas como um todo unificado no processo geral de concretização do direito, através de esquemas de interpretação teleológica - a interpretação corretiva, a redução e a extensão teleológicas. E foi superado, num segundo movimento que este texto retomará com mais vagar no seu sétimo ponto, por uma abertura à teoria da governação: nem o positivismo, nem a hermenêutica, tomados isoladamente, bastam para fundar a função reguladora do direito; ambos requerem, para a sua realização efetiva, uma extensão da reflexividade dos dispositivos institucionais de governação.[71] A superação do formalismo jurídico, por outras palavras, nunca foi apenas hermenêutica: foi sempre também, e desde sempre, governativa e organizativa.

A esta superação vertical do esquema clássico soma-se uma classificação horizontal dos modelos jurisdicionais que o direito comparado e a teoria do direito têm proposto. José Manuel Aroso Linhares distingue dez modelos jurisdicionais sequenciais - do paradigma da aplicação, o *juge bouche de la loi*, ao perfil orgânico-funcional de garantia, o *juge arbitre*, ao modelo normativo-tecnocrático, o *juge entraîneur*, passando pelo juiz sociológico dos consensos político-constitucionais, pelo juiz «centro do sistema», pelo projeto teleológico da jurisdição como engenharia social, pelo juiz da análise económica, pela argumentação institucional combinada, pela relação entre o legislador estratega e o juiz tático, até ao modelo de maior adaptabilidade, o «juiz Hermes», assente num direito líquido reconhecido e experimentado.[72] Esta escala de intensidade criativa importa diretamente para a questão que este texto persegue: não é indiferente automatizar segundo o paradigma do *juge bouche de la loi*, em que a intervenção tecnológica apenas reproduziria um ideal já formalista, ou

70Coelho, Nuno (2024). Ob. cit., p. 157.
71Coelho, Nuno (2024). Ob. cit., pp. 157-160.
72Linhares, J. M. A. (2010). Ob. cit., pp. 450-453.

segundo o do juiz «centro do sistema», em que essa mesma intervenção incidiria sobre um modelo já assumidamente aberto e ponderativo. José Lamego oferece uma taxonomia complementar, mais próxima das escolas da teoria do direito: o modelo lógico-dedutivo, o modelo prático-argumentativo, o modelo material-axiológico - onde situa, precisamente, Arthur Kaufmann e Castanheira Neves - e o modelo do realismo jurídico e dos *Critical Legal Studies*.[73]

É agora possível formular com precisão a equação que este texto persegue. O esquema clássico, os dez modelos de Aroso Linhares e os quatro modelos de Lamego decompõem, cada um ao seu nível de abstração, a mesma linha divisória que a arquitetura da abertura, da moldura e do fechamento já havia traçado, e que o modelo IRAC, nascido noutra tradição jurídica, torna operacionalmente visível: de um lado, operações de recorte mais lógico-classificatório e vinculado à fonte - a determinação da fonte, os elementos gramaticais, lógico-sistemáticos e históricos, o paradigma do *juge bouche de la loi*, o modelo lógico-dedutivo - que correspondem, na grelha do IRAC, às fases de *Issue* e de *Rule*; do outro, operações de recorte valorativo, teleológico e institucionalmente situado - a ponderação das consequências, os elementos teleológicos, o juiz «centro do sistema» ou o «juiz Hermes», o modelo material-axiológico - que correspondem às fases de *Application* e de *Conclusion*. O modelo IRAC não introduz, por isso, uma distinção nova no nosso debate metodológico: comprime, em quatro letras operacionalmente económicas, uma linha divisória que a metodologia jurídica tradicional já vinha traçando, por vias diversas, há várias décadas. É esta genealogia, e não a mera importação de uma sigla anglo-saxónica, que confere ao modelo IRAC a autoridade analítica com que este texto o vai agora utilizar.

O modelo IRAC

É esta arquitetura de abertura, moldura e fechamento que o modelo IRAC - *Issue, Rule, Application, Conclusion* - permite localizar com precisão operacional dentro da própria estrutura da decisão judicial. Trata-se de uma instanciação mais restrita e operativa da mesma problemática, útil precisamente porque decompõe a decisão em fases sucessivas e torna visível a fronteira entre o que a inteligência artificial pode assistir e o que pertence, em exclusivo, ao juiz.

73Lamego, J. (2016). Elementos de metodologia jurídica. Coimbra: Almedina, pp. 175-208.

Fase	Conteúdo	Papel da IA	Responsabilidade
Issue	Identificação da questão jurídica suscitada pelos factos.	Alto potencial: extração de questões, pesquisa inteligente.	IA apoia - juiz valida
Rule	Seleção da norma aplicável: hierarquia, temporalidade, precedência.	Potencial considerável: RAG, jurisprudência dominante.	IA apoia - juiz valida
Application	Aplicação da norma aos factos: interpretação, valores, analogia.	Limite técnico: cláusulas gerais, valoração ética.	Exclusiva do juiz
Conclusion	A decisão fundamentada e a responsabilidade perante as partes.	A IA prevê - não conclui, não fundamenta.	Exclusiva do juiz

Por que a Aplicação e a Conclusão resistem: subsunção e tipo

As fases de *Issue* e de *Rule* - identificação da questão jurídica suscitada pelos factos, seleção da norma aplicável entre hierarquia, temporalidade e precedência - admitem assistência qualificada da inteligência artificial: extração de questões, pesquisa inteligente, geração aumentada por recuperação sobre jurisprudência dominante. Isto não surpreende: são fases em que a operação relevante se aproxima do que Yuk Hui, já referido no segundo ponto deste texto, descrevia como cálculo e otimização recursiva - precisamente o que a máquina humeana executa bem.

As fases de *Application* e de *Conclusion* resistem por uma razão mais precisa do que a mera complexidade técnica. Arthur Kaufmann distinguiu o pensamento subsuntivo, próprio dos conceitos gerais abstratos que definem um objeto através de um número fixo de características fechadas e unívocas - ou se verificam, ou não se verificam - do pensamento tipológico, próprio dos conceitos jurídicos aplicados, que não conhecem o tudo ou nada da subsunção, mas antes o mais ou menos de um tipo que se aproxima ou se afasta de um padrão nunca inteiramente fixado. Este pensamento tipológico é, para Kaufmann, um pensamento

analógico - muito distante, nas suas palavras, do simples pensamento subsuntivo.[74] Podemos formular a mesma ideia em termos mais próximos da dogmática portuguesa: a regra jurídica é um critério material de qualificação, que permite decidir casos concretos precisamente porque contém, em si, uma apreciação valorativa generalizadora sobre uma categoria de situações.[75] Aplicar uma norma a um caso não é, portanto, verificar se um conjunto de propriedades está ou não presente - é exercer, sobre um tipo normativo aberto, um juízo de aproximação que mobiliza precisamente a pré-compreensão de que se falou atrás, e que se inscreve na moldura institucional já descrita.

Não se ignora que sistemas híbridos de inteligência artificial, combinando bases de dados estruturadas com aprendizagem automática, possam produzir resultados formalmente indistinguíveis de uma analogia jurídica bem-sucedida. Mas isso apenas confirma, por outra via, a distinção entre a moldura hermenêutica e a moldura institucional já traçada: o sistema identifica padrões que a doutrina e a jurisprudência humanas previamente decantaram, sem que essa identificação equivalha à validação comunitária e axiológica que confere autoridade à analogia enquanto operação jurídica. É por isso que a responsabilidade pela *Conclusion* - a decisão fundamentada e a responsabilidade perante as partes - permanece exclusiva do juiz: a inteligência artificial prevê, não conclui; e não conclui, precisamente, porque concluir pressupõe o exercício de um juízo tipológico e valorativo sobre um tipo normativo aberto, inscrito numa moldura institucional que nenhum sistema técnico integra.

Todavia, mesmo aqui, não podemos deixar de aqui associar um nível mais próximo e imediato de presença do sujeito decisional no recurso às potencialidades do sistema de inteligência artificial. O juiz projeta-se, num regime metacognitivo, na sua capacidade de formular problemas complexos, escrutinar saídas probabilísticas e integrar fragmentos computacionais nas suas decisões responsáveis. Nesta mesma linha, a ponderação das consequências da decisão, em face dos vários cenários possíveis, o aprimoramento da reflexão crítica, com a destilação de pesquisa ou o desenviesamento cognitivo, podem solidificar e potenciar a ponderação na decisão final, mas não a podem substituir.

74Kaufmann, A. (2002). Ob. cit.

75Ascensão, J. de O. (2001). O direito - Introdução e teoria geral (11.ª ed.). Coimbra: Almedina, p. 481.

Um limite estrutural, não uma lacuna tecnológica

Importa fechar este ponto com uma precisão que evita um mal-entendido comum. O limite aqui descrito - entre *Issue* e *Rule*, de um lado, e *Application* e *Conclusion*, do outro; entre a moldura hermenêutica e a moldura institucional; entre subsunção e tipo - não é uma lacuna tecnológica que sistemas mais sofisticados, treinados com mais dados ou com arquiteturas mais avançadas, um dia colmatarão. É um limite estrutural, decorrente de a inteligência artificial não ser mente, e de a interpretação jurídica em sentido pleno pressupor precisamente aquilo que só uma mente situada, com pré-compreensão e responsabilidade, pode oferecer. Uma caixa negra algorítmica pode tornar-se mais precisa quanto ao resultado sem deixar de ser, quanto ao processo, uma caixa negra - e é estruturalmente incompatível com o dever constitucional de fundamentação que a *Application* e a *Conclusion* do modelo IRAC exigem.

É precisamente porque este limite é estrutural, e não uma questão de maturidade tecnológica, que a formação dos profissionais da justiça não pode ser pensada como uma corrida para acompanhar a evolução das ferramentas. O que está em causa, como o sexto ponto deste texto desenvolverá, é cultivar deliberadamente a pré-compreensão, a competência hermenêutica e o juízo tipológico que nenhuma atualização de software um dia tornará dispensáveis.

VI. IMPLICAÇÕES NA FORMAÇÃO JURÍDICA E JUDICIÁRIA. A CULTURA EPISTEMOLÓGICA DE CARIZ TECNOLÓGICO

O ponto anterior fechou com uma advertência: o limite que separa o que a inteligência artificial pode assistir do que pertence, em exclusivo, ao juiz é estrutural, não uma questão de maturidade tecnológica - e é por isso que a formação dos profissionais da justiça não pode ser pensada como uma corrida para acompanhar a evolução das ferramentas. Este sexto ponto retoma essa advertência e pergunta, em concreto, o que significa cultivar deliberadamente a pré-compreensão, a competência hermenêutica e o juízo tipológico no seio dos sistemas de formação judiciária.

Da literacia digital à cultura epistemológica tecnológica

A literacia digital jurídica - saber distinguir entre ferramentas, conhecer as suas funcionalidades e os seus riscos - é uma competência essencial, mas já não é suficiente perante a evolução galopante a que se assiste no plano tecnológico. O que se exige aos profissionais da justiça é antes uma cultura epistemológica de cariz tecnológico: a compreensão de como a tecnologia e a inteligência artificial influenciam a produção, a interpretação e a aplicação do direito, para além da mera utilização instrumental de ferramentas. Esta exigência não é uma novidade introduzida por este texto; decorre diretamente da tese, já desenvolvida no seu segundo ponto, de que o sistema cognitivo está imerso no mundo social e tecnológico, com ele interage e com ele se potencia, e de que as mutações nas redes de comunicação não deixam de afetar o próprio direito.[76]

Esta mesma conclusão é alcançada, por uma via inteiramente distinta, por um recente grupo de investigação do Alan Turing Institute: Drew Hemment e mais de quarenta coautores sustentam que os sistemas de inteligência artificial atuais enfrentam uma limitação fundamental - são moldados por métricas operacionais estreitas que não refletem a diversidade, a ambiguidade e a riqueza da experiência humana - e propõem que as humanidades, as artes e as ciências sociais qualitativas sejam integradas de forma constitutiva, e não meramente complementar, ao desenvolvimento técnico da inteligência artificial.[77] A convergência é reveladora: tanto a epistemologia jurídica como a investigação técnica sobre inteligência artificial chegam, por caminhos independentes, à mesma conclusão - a profundidade interpretativa não é um acrescento cultural à tecnologia, mas uma condição da sua utilização responsável, e a formação dos profissionais que a utilizam tem de a refletir.

O trabalhador metacognitivo: uma pauta para a formação judiciária

Um recente estudo sobre a mutação do trabalho intelectual na era pós-modelos de linguagem de grande escala oferece uma pauta útil, ainda que desenvolvida fora do direito, para pensar esta formação. Os autores descrevem a transição de um regime do conhecimento, centrado na perícia técnica do especialista, para um regime metacognitivo, em que o valor do trabalho se desloca da execução para a capacidade de formular problemas complexos,

76Coelho, N. (2024). Ob. cit., pp. 186-190, 341-343.
77Hemment, D., Kommers, C., et al. (2025). Doing AI differently: Rethinking the foundations of AI via the humanities [White paper]. Londres: The Alan Turing Institute.

escrutinar saídas probabilísticas e integrar fragmentos computacionais em decisões responsáveis. Neste regime, o profissional exerce quatro papéis: o de *designer* de problemas, que traduz incertezas em questões estruturadas; o de orquestrador de agentes, que define protocolos e critérios de paragem; o de avaliador crítico, que escrutina alucinações e testa hipóteses adversarialmente; e o de integrador de sentido, cuja tarefa - indelegavelmente humana - consiste em sintetizar as saídas algorítmicas em decisões contextualizadas, dotadas de significado ético e de responsabilidade moral.[78]

Este último papel - o de integrador de sentido - é, no fundo, uma outra designação para aquilo que os pontos segundo e quinto deste texto descreveram como pré-compreensão, juízo reflexionante e juízo tipológico. A formação judiciária não parte, por isso, de um vocabulário estranho à sua própria tradição metodológica quando se apropria desta pauta; reconhece antes, num registo empírico e contemporâneo, a mesma distinção estrutural entre o que pode ser assistido e o que exige autoria pessoal.

A urgência desta formação não é especulativa. Um estudo experimental pré-registado com 758 consultores de gestão, conduzido por Fabrizio Dell'Acqua e coautores de Harvard, do MIT e da Warwick Business School, demonstrou a existência de uma fronteira tecnológica irregular: dentro do domínio em que o modelo de inteligência artificial fora treinado, a produtividade aumentou 12,2% e a qualidade melhorou; fora dessa fronteira, a utilização acrítica da mesma ferramenta provocou uma quebra de 19 pontos percentuais na probabilidade de acerto.[79] A fronteira não é visível a olho nu, nem se mantém estável de tarefa para tarefa. Formar um juiz para a era da inteligência artificial não pode, por isso, resumir-se a ensinar a usar as ferramentas disponíveis: tem de ensinar a reconhecer, em cada caso concreto, se este se situa dentro ou fora dessa fronteira — o que é, precisamente, um exercício de discernimento hermenêutico, não de competência técnica.

78Meira, S., Neves, A., Belfort, R., Calegario, F., Garcia, V., & Guimarães, C. (2026). As três inteligências, revisitadas: Tríade sinérgica, redes de agentes e orquestração metacognitiva do trabalho na era pós-LLM. Recife: Porto Digital.

79Dell'Acqua, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2025). Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. Organization Science, 37(2), 403-423.

Modelos e instrumentos de formação: da escola judicial ao ambiente imersivo

Nos países de tradição de *civil law*, entre os quais se contam todos os países de língua portuguesa, a formação judiciária organiza-se predominantemente segundo o modelo da escola judicial: os candidatos, recrutados por concurso público após a licenciatura em direito, frequentam uma formação profissional institucionalizada - por exemplo, o Centro de Estudos Judiciários em Portugal, o Instituto Nacional de Estudos Judiciários em Angola, os Centros de Formação Jurídica e Judiciária de Moçambique e de Timor-Leste, a Escola Nacional de Formação e Aperfeiçoamento de Magistrados no Brasil.[80] Este modelo, ao contrário do estágio profissional próprio da tradição de *common law*, permite uma intervenção curricular deliberada e coordenada - precisamente a condição de possibilidade de qualquer estratégia de cultivo da cultura epistemológica tecnológica em escala.

Essa intervenção não pode, contudo, limitar-se a acrescentar uma disciplina de informática jurídica ou mesmo de literacia digital ao currículo existente. Os métodos didáticos devem evoluir para ambientes de aprendizagem imersivos e experienciais - simulações que reproduzam as complexidades práticas e éticas da integração da inteligência artificial em cenários realistas - e as próprias escolas judiciais devem investir no desenvolvimento de competência interna, formando formadores especializados, para assegurar a atualização contínua dos currículos. O recrutamento deve, do mesmo modo, ser reavaliado para valorizar competências híbridas - rigor jurídico tradicional, discernimento tecnológico e bússola ética - sob pena de a formação contínua ter de compensar, tardiamente, uma lacuna que deveria ser prevenida a montante. Um risco adicional merece nota: o de que a própria transformação digital, mal conduzida, amplie desigualdades já existentes no acesso à tecnologia entre magistrados ou mesmo entre regiões. Combater este défice digital através de programas de literacia fundamental e de infraestruturas equitativas não é uma preocupação lateral relativamente à cultura epistemológica tecnológica; é a sua pré-condição material - não há discernimento hermenêutico sobre uma ferramenta a que não se tem acesso.[81]

80Coelho, N., & Lopes, J. M. (2019). Da formação à cooperação: Guia para a formação e cooperação judiciárias na área penal nos países de língua portuguesa. Lisboa: Camões - Instituto da Cooperação e da Língua, pp. 90-113.

81Coelho, Nuno (2025). Literacia digital jurídica: Uma competência essencial na formação judiciária [Texto de apresentação, IV Conselho Coordenador do Centro de Formação Jurídica e Judiciária]. Maputo: CFJJ.

A experiência comparada oferece, a este propósito, exemplos institucionais concretos do que estes ambientes imersivos podem significar na prática, para além do enunciado do princípio. O Grupo de Trabalho para a Digitalização da *European Judicial Training Network*, constituído em 2023, organiza-se em torno de quatro objetivos - competências digitais, uso ético e legal da inteligência artificial, aprendizagem imersiva e cooperação transnacional - e mantém um Programa de Embaixadores Digitais que projeta a transformação curricular para o plano nacional de cada Estado-Membro.[82] O Ciclo Aprofundado do Digital da *École Nationale de la Magistrature* francesa - um programa anual, com cinco sequências e uma viagem internacional - concretiza, em formato replicável, a exigência de ambientes de aprendizagem experienciais já enunciada.[83] E, fora da tradição de escola judicial que estrutura a formação lusófona, o *Federal Judicial Center* norte-americano condensa a mesma exigência de discernimento hermenêutico em quatro perguntas operativas que todo o juiz deveria colocar perante um sistema de inteligência artificial: como está a ser utilizado para informar a decisão; se o próprio decisor compreende as suas forças, limitações e riscos, incluindo o viés; se a sua aplicação é autêntica, relevante e fiável para a questão em causa; e se a decisão final foi tomada por um algoritmo, por um humano, ou por ambos, de forma documentada e passível de revisão.[84] A convergência entre uma tradição de *civil law* e uma de *common law* nesta mesma exigência de discernimento - e não de mera literacia instrumental - confirma, por via comparada, a tese central deste ponto.

A cooperação dos países de expressão portuguesa como espaço de formação

A formação judiciária atua também como vetor de coesão institucional, fomentando uma cultura judiciária comum que transcende as barreiras corporativas e facilita a cooperação, tanto nacional como internacional. Esta dimensão cooperativa ganha densidade específica no espaço da língua portuguesa: Nuno Coelho e José Mouraz Lopes, num guia de referência para a formação e a cooperação judiciárias na área penal nos países de língua portuguesa, exploraram precisamente o potencial deste espaço, incluindo os desafios tecnológicos que lhe são inerentes, e a possibilidade de uma plataforma de colaboração no

82European Judicial Training Network. (2025). Ob. cit.
83*École Nationale de la Magistrature. (2025). Ob. cit.*
84Federal Judicial Center. (2025). Ob. cit.

âmbito da Conferência dos Ministros da Justiça de Língua Portuguesa.[85] A dinamização da Rede de Escolas de Formação Judiciária dos Países de Língua Oficial Portuguesa permitiria, em particular, a troca sistemática de conhecimentos e de boas práticas entre escolas judiciais que, isoladamente, dificilmente reuniriam massa crítica suficiente para enfrentar os desafios da transformação tecnológica: nenhuma instituição, e nenhum país, pode enfrentar sozinho estes desafios.[86]

Este horizonte cooperativo começa já a ganhar corpo institucional. O projeto Pro Just PALOP-TL, que dá continuidade ao anterior Projeto de Apoio à Consolidação do Estado de Direito, prevê especificamente o reforço dos centros de formação jurídica e judiciária nacionais, a instituição da Rede de Escolas de Formação Jurídica e Judiciária dos Países de Língua Portuguesa - que deverá integrar também o recém-criado Centro de Estudos Jurídicos e Judiciários de Cabo Verde - e a disponibilização de formação especializada em justiça digital. O modelo de intercâmbio horizontal entre pares que a *African Judicial Exchange Network* pratica e institucionalizou em estatutos próprios aprovados em Zanzibar em 2024, oferece aqui um paradigma comparativo relevante: uma cooperação entre instituições de maturidade tecnológica desigual, sustentada não na uniformização mas na troca sistemática de boas práticas - precisamente o modelo que a rede lusófona, dadas as assimetrias de recursos entre Portugal, o Brasil e os demais Estados-membros da CPLP, terá de negociar.[87]

Um estudo aprofundado sobre a capacitação judiciária nos países de língua oficial portuguesa - Angola, Brasil, Cabo Verde, Guiné-Bissau, Moçambique, Portugal, São Tomé e Príncipe e Timor-Leste - que mapeasse as necessidades e competências existentes e futuras dos profissionais da justiça, tendo em conta os diferentes níveis de desenvolvimento tecnológico de cada país, permitiria, por outro lado, transformar esta cultura judiciária comum, hoje ainda em construção, numa verdadeira área de formação judiciária de língua portuguesa, coesa e interligada.

85Coelho, N., & Lopes, J. M. (2019). Ob. cit., pp. 117-130.
86Coelho, N. (2025). Ob. cit..
87Cf. African Judicial Training Network (AJTN), Constitution and inaugural documents, Zanzibar, 2024; CEELI Institute, African Judicial Exchange Network.

Uma cultura, e não só um currículo

Fica, no entanto, uma advertência final. Uma cultura epistemológica não se instaura por decreto curricular, nem se esgota num programa de formação, por mais bem desenhado que seja. Peter Senge descreveu, a propósito das organizações que aprendem, a diferença entre uma organização que acumula conhecimento técnico e uma organização que desenvolve, coletivamente, a capacidade de questionar os seus próprios pressupostos.[88] É esta segunda capacidade - reflexiva, coletiva, permanentemente inacabada - que os sistemas de formação judiciária têm de cultivar, e não só a primeira. Encarar os tribunais e o sistema judicial no seu conjunto como organizações de aprendizagem, nesta aceção, desloca as estruturas de formação e de capacitação para o eixo central das preocupações institucionais, e não para a sua periferia administrativa.[89] Sem essa capacidade cultivada, todos os instrumentos descritos neste ponto - os ambientes imersivos, os currículos revistos, as plataformas de cooperação lusófona - arriscam-se a produzir profissionais tecnicamente competentes, mas hermeneuticamente desarmados perante o único risco que verdadeiramente importa prevenir: o de confundir a fluência de uma resposta com a justeza de uma decisão.

É esta mesma preocupação - a de que a formação, por si só, não basta sem um enquadramento institucional que a sustente e a torne exigível - que conduz este texto, no seu sétimo e último ponto, aos desafios da governação, da coordenação e da cooperação que a integração da inteligência artificial na justiça coloca aos Estados e às instituições.

VII. CONCLUSÃO. DESAFIOS DE GOVERNAÇÃO, COORDENAÇÃO E COOPERAÇÃO

O ponto anterior fechou com uma advertência: a formação, por si só, não basta sem um enquadramento institucional que a sustente e a torne exigível. Este sétimo e último ponto ocupa-se desse enquadramento, nos seus três planos - a governação interna dos tribunais, a coordenação entre os diferentes níveis normativos que já se foram cruzando ao longo deste texto, e a cooperação internacional que a natureza global da tecnologia torna inevitável - antes de regressar, no seu termo, ao ponto de partida.

88Senge, P. M. (1990). The fifth discipline: The art and practice of the learning organization. Nova Iorque: Doubleday/Currency.
89Coelho, N. (2025). Ob. cit.

Governação: a soberania do código posta à prova

A governação tecnológica dos tribunais deixou de ser uma preocupação abstrata para se tornar uma experiência concreta e recente. Em outubro de 2025, a migração dos sistemas da jurisdição administrativa e fiscal portuguesa para a plataforma unificada dos tribunais judiciais precipitou aquilo que se veio a designar por Outubro Negro: cerca de 140 mil processos relacionados com imigração e asilo ficaram invisíveis ou foram incorretamente encaminhados, com a distribuição algorítmica a atribuir processos a magistrados errados - numa violação, ainda que não intencional, do princípio do juiz natural.[90] O incidente repetiu, uma década volvida, o colapso de setembro de 2014 que paralisou os tribunais portugueses durante 44 dias por ocasião da migração para o novo mapa judiciário, e confirma, por via empírica, e não só teórica, o que este texto sustentou desde o seu quarto ponto: a governação do Nível I - a otimização da gestão, aparentemente a mais distante do ato de julgar - não é menos crítica do que a governação dos níveis mais próximos da decisão; é, na verdade, a sua pré-condição.

A soberania do código - a capacidade dos tribunais de controlarem a sua própria infraestrutura tecnológica, em vez de a delegarem, sem resguardo, em fornecedores privados - não é, por isso, um enunciado retórico. É a condição de possibilidade de todos os princípios de supervisão descritos no quarto ponto deste texto: sem controlo sobre o código, não há *Human ON-the-loop* que resista à *shadow AI*, nem *Human IN Command* que resista à opacidade de uma caixa negra fornecida por terceiros.

Seis princípios recorrem, com pequenas variações terminológicas, na literatura comparada sobre governação de sistemas de inteligência artificial em contexto organizacional, e sintetizam o que a soberania do código exige em concreto: a clareza e a compreensibilidade da atuação dos sistemas para todos os interessados; a transparência quanto à lógica de decisão e às suas limitações; a resiliência técnica e a segurança perante cenários inesperados; o uso responsável de dados e a privacidade; a responsabilização e a atribuição clara de funções ao longo de todo o ciclo de vida do sistema; e a centralidade humana e a supervisão, que garantem que a tecnologia amplia, e não substitui, as capacidades do julgador.[91] Estes

90Coelho, Nuno (2025). A soberania do código e a autonomia tecnológica: Os imperativos da justiça na era do digital e da IA. Lisboa: Observador.

91Cf. National Institute of Standards and Technology. (2023). AI risk management framework. Washington, DC: U.S. Department of Commerce, e a literatura de governação organizacional de IA nele inspirada.

princípios não são compartimentos estanques: a transparência é condição da responsabilização, e o uso responsável de dados condiciona diretamente a mitigação do viés - pelo que uma governação tecnológica eficaz na justiça exige uma abordagem holística, e não a sua justaposição posterior a um sistema já adquirido e já em produção.

Em Portugal, esta exigência de governação articula-se, hoje, num quadro normativo já com alguma densidade, sem perder as dimensões multinível da regulação europeia: as Recomendações do Conselho Superior da Magistratura, já referidas, para a jurisdição comum, e a Carta Ética do Conselho Superior dos Tribunais Administrativos e Fiscais, de 25 de março de 2025, para a jurisdição administrativa e fiscal - esta última alertando, num registo particularmente lúcido, para o risco de uma «*nova normatividade*», uma uniformização das decisões assente em estatísticas e não numa análise caso a caso da aplicação da lei.[92] Ambos os instrumentos se articulam, por sua vez, com a Carta Ética da CEPEJ e com o Regulamento europeu sobre inteligência artificial, já citados no quarto ponto deste texto.

Este quadro regulatório europeu, adotado através do Regulamento (UE) 2024/1689, que classifica os sistemas de inteligência artificial aplicados à administração da justiça como sistemas de risco elevado,[93] sofreu entretanto uma reforma de vulto: o Regulamento (UE) 2026/1744, de 8 de julho de 2026 - o «Digital Omnibus em matéria de IA» - adia para 2 de dezembro de 2027 a aplicação das obrigações relativas aos sistemas de risco elevado autónomos do Anexo III, entre os quais se contam os sistemas judiciais, e para 2 de agosto de 2028 a aplicação das obrigações relativas aos sistemas integrados noutros produtos, sem ceder, todavia, um único dia no prazo de 2 de dezembro de 2026 fixado para a nova proibição absoluta de sistemas que gerem conteúdo íntimo não consensual ou material de abuso sexual de menores, e concentrando a fiscalização dos sistemas assentes em modelos de finalidade geral no Gabinete Europeu da IA.[94] O adiamento do calendário vinculativo não diminui, porém, o alcance das salvaguardas já assumidas voluntariamente pelos tribunais portugueses e europeus: as Recomendações do Conselho Superior da Magistratura e a Carta Ética do

92Conselho Superior dos Tribunais Administrativos e Fiscais. (2025). Carta Ética para a Utilização da Inteligência Artificial nos Tribunais Administrativos e Fiscais [Deliberação de 25 de março de 2025]. Lisboa: CSTAF.

93Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, Anexo III, ponto 8, alínea a).

94Regulamento (UE) 2026/1744 do Parlamento Europeu e do Conselho, de 8 de julho de 2026 (Digital Omnibus em matéria de IA). Jornal Oficial da União Europeia, L 2026/1744, 24.7.2026.

Conselho Superior dos Tribunais Administrativos e Fiscais, ambas anteriores a esta reforma, continuam já hoje a antecipar, por via de *soft law* interna, o que o Regulamento só tornará juridicamente exigível dentro de um ou dois anos - confirmando, uma vez mais, que a governação tecnológica dos tribunais não pode esperar pelos prazos legais para começar a proteger o que estes prazos apenas mais tarde virão a impor.

Esta arquitetura normativa não esgota, todavia, o universo de atores cuja coordenação a governação tecnológica dos tribunais reclama: o sistema de formação e capacitação em que se sustenta convoca, para além do poder judicial, o legislador, o Ministério da Justiça, o Ministério Público e as representações das diversas profissões da justiça, cada um com responsabilidades próprias e complementares.[95]

Coordenação: entre o consentimento e a decisão automatizada

Se a governação regula o que se passa dentro de cada tribunal, a coordenação tem de responder a uma pergunta mais ampla e mais perturbadora: até onde pode a decisão automatizada avançar, e com que condições? Sir Geoffrey Vos, Master of the Rolls de Inglaterra e País de Gales, formulou esta pergunta, numa conferência proferida em Sydney em agosto deste ano, através de uma distinção que este texto subscreve: uma coisa são decisões automatizadas consentidas pelas partes - os litígios comerciais de baixo valor já hoje resolvidos por arbitragem assistida por inteligência artificial, com supervisão humana, junto da *American Arbitration Association* - outra, bem diferente, é impor decisões automatizadas a partes que a elas não consentiram. Nesta segunda hipótese, Vos sustenta, e este texto acompanha-o, que dificilmente um sistema de inteligência artificial poderá alguma vez constituir um «*tribunal independente e imparcial*», na aceção do artigo 6.º da Convenção Europeia dos Direitos do Homem ou do artigo 14.º do Pacto Internacional sobre os Direitos Civis e Políticos.[96]

Vos acrescenta uma observação que este texto não havia ainda considerado, e que merece registo pela sua originalidade: a inteligência artificial processa dados e alcança conclusões em segundos, mas os seres humanos são muito mais lentos do que as máquinas a processar,

95Coelho, N. (2025). Ob. cit.

96Vos, G. (2026, 18 de agosto). The implications of the development of artificial intelligence on the judiciary and the legal profession [John Lehane Memorial Lecture]. Supreme Court of New South Wales, Sydney, Austrália.

compreender e aceitar essas mesmas conclusões - levam dias, ou meses. Se a inteligência artificial decide questões que afetam pessoas e negócios num instante, não é razoável esperar que as pessoas se conformem com essas decisões no mesmo intervalo de tempo.[97] Esta assimetria temporal entre a velocidade da máquina e o tempo de legitimação humana é, no fundo, uma reformulação, num registo empírico e institucional, da distinção entre processo e resultado que o quinto ponto deste texto explorou a propósito da hermenêutica: uma decisão pode ser produzida em segundos, mas a moldura institucional que lhe confere autoridade - a aceitação, a legitimação, o assentimento da comunidade jurídica - não se deixa comprimir ao mesmo ritmo. É esta discrepância, e não só a discrepância técnica, que a coordenação normativa tem de gerir.

O artigo 14.º do Regulamento europeu sobre inteligência artificial, já referido, resolve este problema de coordenação através de uma fórmula prudente: qualifica os sistemas judiciais como sistemas de risco elevado, que têm de ser efetivamente supervisionados por pessoas singulares, sem proibir, contudo, que a inteligência artificial aconselhe o juiz. A linha entre o aconselhamento e a decisão automatizada permanece, por isso, aberta - e é exatamente aí, na zona cinzenta entre os níveis II e III da tipologia proposta no quarto ponto deste texto, que a coordenação normativa mais se torna necessária, e mais difícil de garantir na prática quotidiana dos tribunais.

Esta dificuldade não é meramente teórica. Uma revisão de âmbito recente, encomendada pela Nuffield Foundation e incidente sobre a justiça cível e familiar inglesa e galesa, identificou quarenta e cinco ferramentas de inteligência artificial já em uso ou em desenvolvimento nos tribunais, das quais apenas sete dispunham de avaliação independente publicada - e mesmo estas concentradas em ganhos de produtividade e de eficiência, e não nos efeitos sobre a equidade, os resultados jurídicos ou a experiência de quem recorre à justiça.[98] A conclusão da autora é direta: a base de evidência disponível não permite ainda responder à pergunta central de saber se a inteligência artificial melhora, na prática, o acesso à justiça — o que torna a avaliação independente e robusta destas ferramentas uma prioridade urgente, e não uma diligência a realizar depois de instaladas. É este défice de verificação, e não apenas a zona cinzenta normativa já identificada, que a coordenação entre os níveis II e III da tipologia

97Vos, G. (2026). Ob. cit.
98Sargeant, H. (2026). Ob. cit.

proposta no quarto ponto deste texto terá também de resolver - e que a magistratura portuguesa, ainda numa fase mais incipiente de adoção destas ferramentas, faria bem em antecipar, em vez de se limitar a herdar.

Cooperação: a responsabilidade civil como fronteira comum

Um terceiro plano, ainda pouco explorado pela doutrina de expressão portuguesa e que este texto identifica, no seu termo, como uma fronteira de investigação futura, é o da responsabilidade civil por danos causados pela inteligência artificial. Sir Geoffrey Vos alerta que, decorridos quase quatro anos sobre o lançamento público de sistemas de inteligência artificial generativa, ainda existe uma incerteza jurídica considerável sobre quem responde quando esses sistemas causam prejuízo - os fornecedores de modelos de linguagem, os profissionais que os utilizam, ou ambos.[99] Na União Europeia, esta incerteza tem sido gerida através da diretiva revista relativa à responsabilidade decorrente de produtos defeituosos, depois de a Comissão Europeia ter retirado, em fevereiro de 2025, a proposta de uma diretiva especificamente dedicada à responsabilidade por inteligência artificial; no Reino Unido, o *UK Jurisdiction Taskforce*, presidido pelo próprio Vos, publicou em julho de 2026 um parecer jurídico sobre a responsabilidade por danos causados por inteligência artificial ao abrigo do direito inglês.

Esta incerteza não é um problema exclusivamente europeu ou anglófono; é, precisamente por isso, um problema que reclama cooperação internacional. O *Standing International Forum of Commercial Courts*, de que os tribunais superiores de dezenas de países são membros, tem procurado construir princípios de consenso alargado sobre o que é e o que não é aceitável no uso da inteligência artificial em decisões comerciais e arbitrais.[100] Este tipo de arquitetura de consenso alargado não carece de precedente institucional: já em 2017, a *International Organization for Judicial Training*, reunindo então cento e vinte e nove instituições de formação de setenta e nove países, aprovara por unanimidade uma declaração de princípios comuns sobre formação judicial - prova de que a construção de consenso internacional sobre padrões da função judicial, mesmo entre culturas jurídicas muito diversas, é possível quando existe vontade institucional para a sustentar.[101] Não existe, ainda, um fórum equivalente no

99Vos, G. (2026). Ob. cit.
100Vos, G. (2026). Ob. cit.
101International Organization for Judicial Training. (2017). Ob. cit.

espaço lusófono - mas a arquitetura que o sexto ponto deste texto descreveu a propósito da formação judiciária, com a Conferência dos Ministros da Justiça de Língua Portuguesa como plataforma emergente de cooperação, oferece precisamente o embrião institucional de que essa resposta poderia partir. A soberania algorítmica, como sublinha um recente estudo sobre a transformação do trabalho na era pós-modelos de linguagem, deve ser tratada como um bem público estratégico, comparável à energia, à saúde e à educação - e bens públicos estratégicos não se protegem isoladamente.[102] Nenhuma instituição, e nenhum país, pode enfrentar isoladamente os desafios que a transformação tecnológica da justiça coloca: a cooperação em rede não é uma opção entre outras, é um imperativo estratégico.[103]

Regresso ao Eu Múltiplo

Este texto partiu, na sua introdução, de uma exposição sobre o eu como múltiplo: a ideia de que a identidade não é uma essência fixa, mas um movimento contínuo de composição - gestos, memórias e influências que se cruzam sem nunca se estabilizarem de vez.[104] Sete pontos depois, é possível regressar a essa imagem com mais do que uma intuição literária: com um argumento. A atividade jurisdicional revelou-se, ao longo deste percurso, tão pluralmente composta quanto o eu da exposição do MACAM - otimização da gestão, apoio à decisão, reflexão crítica; *issue, rule, application, conclusion*; moldura hermenêutica, moldura institucional; literacia digital, cultura epistemológica. Nenhum destes níveis, por si só, esgota o que significa julgar.

Mas, e este é o ponto que todo o texto procurou defender desde o segundo ponto até este último, esta pluralidade não dissolve a unidade do julgador, tal como a pluralidade do eu, na leitura que a exposição propunha, não dissolvia a identidade, antes a constituía enquanto diálogo contínuo entre o visível e o invisível, entre o ser e o tornar-se.[105] A jurisdição permanece uma só, mesmo multiplicada pelos seus níveis de intervenção tecnológica, precisamente porque um núcleo resiste, em cada um desses níveis, à dissolução: o juízo tipológico, a pré-compreensão, a responsabilidade pessoal do julgador perante o caso concreto e perante a comunidade. É este núcleo que a governação, a coordenação e a

102Meira, S., Neves, A., Belfort, R., Calegario, F., Garcia, V., & Guimarães, C. (2026). Ob. cit.
103Coelho, N. (2025). Ob. cit.
104Quintela, C. (2025). Ob. cit.
105Quintela, C. (2025). Ob. cit.

cooperação aqui descritas existem para proteger - não para o substituir, nem para o dispensar, mas para lhe permitir persistir, múltiplo e uno, numa era que já não lhe permite ser simples.

Fecha-se este texto, por isso, não com certezas fechadas, mas com quatro perguntas que se espera possam alimentar o debate que se lhe segue. Que deveres de transparência devem recair sobre os advogados quanto ao uso de inteligência artificial generativa na elaboração de peças processuais, e como articular esses deveres com o dever de menção que recai sobre o juiz? Como evitar que a adoção de sistemas de inteligência artificial na gestão processual dos tribunais introduza, a montante, vieses sistémicos que condicionem a informação que chega ao gabinete do juiz sem que este conheça a sua origem? É possível desenhar formação contínua que previna o *deskilling*, mantendo ativas as competências hermenêuticas nucleares num ambiente de assistência algorítmica generalizada? E, por fim, a pergunta que talvez condense todas as outras: que juiz, ou que eu múltiplo do juiz, queremos que a tecnologia ajude a compor, e que responsabilidade estamos dispostos a assumir, coletiva e institucionalmente, por essa composição? Não uma justiça substituída: uma justiça aumentada, em que o algoritmo liberta o juiz - desde que o juiz e as instituições que o garantem, não se deixem libertar da sua própria responsabilidade.

REFERÊNCIAS BIBLIOGRÁFICAS E DOCUMENTAIS

ALBERTO, N. (2024). *Algoritmos e big data a partir do sistema de justiça criminal português: Contributos para uma justiça automatizada*. Coimbra: Almedina.

ASCENSÃO, J. de O. (2001). *O direito - Introdução e teoria geral* (11.ª ed.). Coimbra: Almedina.

BAPTISTA MACHADO, J. (2017). *Introdução ao direito e ao discurso legitimador*. Coimbra: Almedina. (1.ª ed. 1983.)

BEN-MICHAEL, E., Greiner, D. J., Huang, M., Imai, K., Jiang, Z., & Shin, S. (2025). "Does AI help humans make better decisions? A statistical evaluation framework for experimental and observational studies". *Proceedings of the National Academy of Sciences*, 122(38), e2505106122.

CAMPOS, Ricardo (2022). *Metamorfoses do direito global: Sobre a interação entre direito, tempo e tecnologia*. São Paulo: Contracorrente.

CANOTILHO, J. J. G. (2003). *Direito constitucional e teoria da constituição* (7.ª ed.). Coimbra: Almedina.

CASTANHEIRA NEVES, A. (1993). *Metodologia jurídica: Problemas fundamentais*. Coimbra: Coimbra Editora.

CEPEJ - Comissão Europeia para a Eficiência da Justiça. (2018). *Carta Ética Europeia sobre o Uso da Inteligência Artificial nos Sistemas Judiciais e no seu Ambiente*. Estrasburgo: Conselho da Europa.

CLARK, A., & Chalmers, D. (1998). "The extended mind". *Analysis*, 58(1), 7–19.

COECKELBERGH, M. (2022). "Democracy, epistemic agency, and AI: Political epistemology in times of artificial intelligence". *AI and Ethics*. https://doi.org/10.1007/s43681-022-00239-4

COELHO, Nuno (2024). *Organização do sistema judicial e jurisdição: Para uma sistemática alargada da realização do direito pelo juiz*. Coimbra: Almedina.

COELHO, Nuno (2025). "A soberania do código e a autonomia tecnológica: Os imperativos da justiça na era do digital e da IA". Artigo de Opinião, Observador, Lisboa: Observador.

COELHO, Nuno (2025). Literacia digital jurídica: Uma competência essencial na formação judiciária [Texto de apresentação, IV Conselho Coordenador do Centro de Formação Jurídica e Judiciária, Moçambique], Mimeo.

COELHO, Nuno (2026). IA na fase escrita, nas audiências e na decisão jurisdicional [Texto de apresentação, 3.º Painel, X Conferência Diálogos Magistrados-Advogados: Justiça e Inteligência Artificial]. Mimeo.

COELHO, Nuno, & Lopes, José Mouraz (2019). *Da formação à cooperação: Guia para a formação e cooperação judiciárias na área penal nos países de língua portuguesa*. Lisboa: Camões - Instituto da Cooperação e da Língua.

COFONE, I. (2021). "AI and judicial decision-making". Em F. Martin-Bariteau & T. Scassa (Eds.), *Artificial intelligence and the law in Canada*. Toronto: LexisNexis.

CONSELHO SUPERIOR DA MAGISTRATURA DE PORTUGAL. (2026). *Recomendações para o Uso de Inteligência Artificial na Atividade Judicial* [GAIA 2025/GAVPM/5469]. Lisboa: CSM.

CONSELHO SUPERIOR DOS TRIBUNAIS ADMINISTRATIVOS E FISCAIS. (2025). *Carta Ética para a Utilização da Inteligência Artificial nos Tribunais Administrativos e Fiscais* [Deliberação de 25 de março de 2025]. Lisboa: CSTAF.

CONTINI, F. (2020). "Artificial intelligence and the transformation of humans, law and technology interactions in judicial proceedings". *Law, Technology and Humans*, 2(1), 4–18.

DELL'ACQUA, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2025). "Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality". *Organization Science*, 37(2), 403-423.

DESMARAIS, S. L., Zottola, S. A., Duhart Clarke, S. E., & Lowder, E. M. (2021). "Predictive validity of pretrial risk assessments: A systematic review of the literature". *Criminal Justice and Behavior*, 48(4), 398-420.

DWORKIN, R. (1978). *Taking rights seriously*. Cambridge, MA: Harvard University Press.

DYEVRE, A., & Shahvaroughi, A. (2026). *Man and machine: Artificial intelligence and judicial decision making* (preprint). Lovaina: Leuven Centre for Empirical Jurisprudence.

ÉCOLE NATIONALE DE LA MAGISTRATURE. (2025). *Catalogue de formation continue* 2025. Paris: ENM.

EUROPEAN JUDICIAL TRAINING NETWORK. (2025). *Digitalisation portfolio and Working Group on Digitalisation*. Bruxelas: EJTN.

FEDERAL JUDICIAL CENTER. (2025). *An introduction to artificial intelligence for federal judges*. Washington, DC: FJC.

FUCHS, T. (2021). *In defense of the human being: Foundational questions of an embodied anthropology*. Oxford: Oxford University Press.

GADAMER, H.-G. (1995). *Truth and method* (2.ª ed. rev. inglesa da 5.ª ed. alemã de 1986). Nova Iorque: The Continuum Publishing Company.

GARAPON, A. (1996). *Le gardien des promesses: Justice et démocratie*. Paris: Odile Jacob.

GARAPON, A., Perdriolle, S., Bernabé, B., & Kadri, C. (2013). *La prudence et l'autorité: L'office du juge au XXIe siècle* [Rapport]. Paris: Institut des Hautes Études sur la Justice (IHEJ).

GODDARD, K., Roudsari, A., & Wyatt, J. C. (2012). "Automation bias: A systematic review of frequency, effect mediators, and mitigators". *Journal of the American Medical Informatics Association*, 19(1), 121-127.

GUNKEL, D. J. (2012). *The machine question: Critical perspectives on AI, robots, and ethics*. Cambridge, MA: MIT Press.

HABERMAS, J. (1970). *Toward a rational society: Student protest, science, and politics* (J. J. Shapiro, Trad.). Boston: Beacon Press.

HEIDEGGER, M. (1977). *The question concerning technology and other essays* (W. Lovitt, Trad.). Nova Iorque: Garland Publishing.

HEMMENT, D., Kommers, C., et al. (2025). *Doing AI differently: Rethinking the foundations of AI via the humanities* [White paper]. Londres: The Alan Turing Institute.

HIRSCH BALLIN, E., van der Schyff, G., & Stremler, M. (Eds.). (2020). *European Yearbook of Constitutional Law 2019: The Role of the Judiciary in Contemporary Constitutional Systems*. Haia: T.M.C. Asser Press.

HUI, Y. (2026). *Kant machine: Critical philosophy after AI*. Londres: Bloomsbury Academic.

INTERNATIONAL ORGANIZATION FOR JUDICIAL TRAINING. (2017). *Declaration of principles of judicial education*. Nova Deli: IOJT.

KAHL, P. (2025). *The epistemic architecture of power: How knowledge control sustains authority in social structures* (2.ª ed.). Londres: Lex et Ratio Ltd. https://doi.org/10.13140/RG.2.2.14628.54402

KALLINIKOS, J. (2005). "The order of technology: Complexity and control in a connected world". LSE Research Online.

KALLINIKOS, J. (2011). *Governing through technology: Information artefacts and social practice*. Hampshire/Nova Iorque: Palgrave Macmillan.

KANT, I. (2007). *Critique of judgment* (J. C. Meredith & N. Walker, Trad.). Oxford: Oxford University Press.

KAUFMANN, A. (2002). "Filosofia do direito, teoria do direito, dogmática jurídica". Em A. Kaufmann & W. Hassemer (Orgs.), *Introdução à filosofia do direito e à teoria do direito contemporâneas* (pp. 57-161). Lisboa: Fundação Calouste Gulbenkian.

KELSEN, H. (2005). *Teoria pura do direito* (J. B. Machado, Trad., 7.ª ed. alemã). Coimbra: Almedina.

KLEINBERG, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). "Human decisions and machine predictions". *The Quarterly Journal of Economics*, 133(1), 237-293.

LADEUR, K.-H. (2007). "The role of contracts and networks in public governance: The importance of the "social epistemology" of decision making". *Indiana Journal of Global Legal Studies*, 14(2), 329–351.

LADEUR, K.-H. (2017). "The transformation of legal semantics in postmodern society". *Revista de Informação Legislativa*, 54(215), 13–26.

LAMEGO, J. (1990). *Hermenêutica e jurisprudência: Análise de uma "recepção"*. Lisboa: Fragmentos.

LAMEGO, J. (2016). *Elementos de metodologia jurídica*. Coimbra: Almedina.

LANZARA, G. F. (2009). "Building digital institutions: ICT and the rise of assemblages in government". Em F. Contini & G. F. Lanzara (Eds.), *ICT and innovation in the public sector* (pp. 9–48). Nova Iorque: Palgrave MacMillan.

LANZARA, G. F. (2014). "The circulation of agency in judicial proceedings: Designing for interoperability and complexity". Em F. Contini & G. F. Lanzara (Eds.), *The circulation of agency in e-justice*. Dordrecht/Heidelberga/Nova Iorque/Londres: Springer.

LAWRENCE, N. D. (2024). *The atomic human: Understanding ourselves in the age of AI*. Londres: Allen Lane.

LARENZ, K. (1989). *Metodologia da ciência do direito* (5.ª ed. rev., J. Lamego, Trad.). Lisboa: Fundação Calouste Gulbenkian.

LEFORT, C. (1988). *Democracy and political theory* (D. Macey, Trad.). Cambridge: Polity Press.

LINHARES, J. M. A. (2010). "Jurisdição, diferendo e «área aberta»: A caminho de uma teoria do direito como moldura". Em *Ars Ivdicandi — Estudos em Homenagem ao Prof. Doutor Jorge de Figueiredo Dias*, Vol. IV (pp. 443–477). Coimbra: Coimbra Editora.

LINNA, E., & Linna, T. (2026). "Challenges for generative AI in legal reasoning". *Discover Artificial Intelligence*, 6, artigo 170.

LUHMANN, N. (1993a). *Risk: A sociological theory*. Berlim/Nova Iorque: Walter de Gruyter.

LUHMANN, N. (2004). *Law as a social system* (K. A. Ziegert, Trad.). Oxford: Oxford University Press.

MAGGI, B. (1984). "Teoria dell'organizzazione e sociologia del diritto". Em U. Scarpelli & V. Tomeo (Eds.), *Società, norme e valori: Studi in onore di Renato Treves* (pp. 285–307). Milão: Giuffré.

MEIRA, S., Neves, A., Belfort, R., Calegario, F., Garcia, V., & Guimarães, C. (2026). *As três inteligências, revisitadas: Tríade sinérgica, redes de agentes e orquestração metacognitiva do trabalho na era pós-LLM*. Recife: Porto Digital.

MUMFORD, L. (1934). *Technics and civilization*. Nova Iorque: Harcourt, Brace and Company.

NEVES, M. (1994). *A Constitucionalização Simbólica*. São Paulo: Editora Acadêmica.

NOTI, G., Donahue, K., Kleinberg, J., & Oren, S. (2025). *AI-assisted decision making with human learning*. Proceedings of the 26th ACM Conference on Economics and Computation (EC '25), 38-39.

ONGARO, G., Hardman, D., & Deschenaux, I. (2022). "Why the extended mind is nothing special but is central". *Phenomenology and the Cognitive Sciences*. 23(4), 841–863. https://doi.org/10.1007/s11097-022-09827-5

OST, F., & van de Kerchove, M. (2002). *De la pyramide au réseau? Pour une théorie dialectique du droit*. Bruxelas: Publications des Facultés universitaires Saint-Louis.

PEREIRA, A. L. D. (2008). "Lex informática, ius ex machina e justiça artificial". Em J. de Figueiredo Dias, J. J. Gomes Canotilho & J. de Faria Costa (Orgs.), *Ars Iudicandi — Estudos em Homenagem ao Prof. Doutor António Castanheira Neves*, Vol. I (pp. 817–886). Coimbra: Coimbra Editora.

QUATTROCIOCCHI, W., Capraro, V., & Perc, M. (2025). "Epistemological fault lines between human and artificial intelligence". *arXiv:2512.19466*.

QUINTELA, C. (2025). O eu como múltiplo [Catálogo de exposição]. Lisboa: MACAM – Museu de Arte Contemporânea Armando Martins.

REGULAMENTO (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024 (Regulamento Europeu sobre Inteligência Artificial). Jornal Oficial da União Europeia, L 2024/1689.

REGULAMENTO (UE) 2026/1744 do Parlamento Europeu e do Conselho, de 8 de julho de 2026, que altera os Regulamentos (UE) 2024/1689, (UE) 2018/1139 e (UE) 2023/1230 no que respeita à simplificação da aplicação das regras harmonizadas em matéria de inteligência artificial (Digital Omnibus em matéria de IA). Jornal Oficial da União Europeia, L 2026/1744.

RODRIGUES, A. M. (Coord.). (2020). *Inteligência artificial no direito penal*. Coimbra: Almedina.

SARGEANT, H. (2026). *Artificial intelligence & justice: An evidence scoping review*. Londres: Nuffield Foundation.

SARTOR, G., & Branting, L. K. (Eds.). (1998). *Judicial applications of artificial intelligence*. Dordrecht: Kluwer Academic Publishers.

SENGE, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. Nova Iorque: Doubleday/Currency.

SHAW, S. D., & Nave, G. (2026). *Thinking—fast, slow, and artificial: How AI is reshaping human reasoning and the rise of cognitive surrender* (working paper). Filadélfia: The Wharton School, University of Pennsylvania.

SILVA, R. M. M. T. da (2025). *Interpretação artificial: Sobre a possibilidade de os novos sistemas de inteligência artificial fazerem interpretação jurídica*. Coimbra: Almedina.

SIMONDON, G. (1958). *Du mode d'existence des objets techniques*. Paris: Aubier-Montaigne.

SMART, P. (2017). "Extended cognition and the internet: A review of current issues and controversies". *Philosophy & Technology*, 30, 357–390.

SOLOMONOFF, R. (1964). "A formal theory of inductive inference (Parts I & II)". *Information and Control*, 7(1–2), 1–22, 224–254.

SUNSTEIN, C. R. (2001). "Of artificial intelligence and legal reasoning". *Public Law & Legal Theory*, University of Chicago Law School, Working Paper n.º 18.

TARUFFO, M. (1998). "Judicial decisions and artificial intelligence". *Artificial Intelligence and Law*, 6(2-4), 311-324.

TEUBNER, G. (2006). "The anonymous matrix: Human rights violations by "private" transnational actors". *Modern Law Review*, 69(3), 327–346.

TEUBNER, G., & Collins, H. (2011). *Networks as connected contracts*. Oxford/Portland: Hart Publishing.

VALENTE, M. M. G. (Coord.). (2024). *Ciências Criminais e Inteligência Artificial*. Coimbra. Almedina.

VESTING, T. (2004). "The autonomy of law and the formation of network standards". *German Law Journal*, 5(6), 639–668.

VESTING, T. (2017). "Modern law and the crisis of common knowledge". *Revista de Informação Legislativa*, 54(214), 11–29.

VESTING, T. (2018). *Legal Theory and the Media of Law*. Cheltenham/Northampton: Edward Elgar Publishing.

VILLAS BOAS, M. (2024). *O uso ético e estratégico da inteligência artificial generativa no judiciário: O método dialógico multirresposta*. São Paulo: EACH/USP; COPEDEM; ESMAT, Mimeo.

VOLOKH, E. (2019). "Chief Justice robots". *Duke Law Journal*, 68(6), 1135-1192.

VOS, G. (2026, 18 de agosto). "The implications of the development of artificial intelligence on the judiciary and the legal profession" [John Lehane Memorial Lecture]. Supreme Court of New South Wales, Sydney, Austrália.

WAGNER, A., & Marusek, S. (2026). "Reimagining law through the rhizome — Fluidity, fracture, and forward movement" [Editorial]. *Undecidabilities and Law: The Coimbra Journal for Legal Studies*, (6), 7–19. https://doi.org/10.14195/2184-9781_6_0

WIENER, N. (1961). *Cybernetics: Or control and communication in the animal and the machine* (2.ª ed.). Cambridge, MA: MIT Press.

ZALNIERIUTE, M., & Limante, A. (Eds.). (2026). *The Cambridge handbook of AI and technologies in courts*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009744225



INTELIGÊNCIA ARTIFICIAL (IA) E JUÍZES - PISTAS PARA O FUTURO

JOSÉ JOAQUIM FERNANDES OLIVEIRA MARTINS♦

♦ Juiz de Direito, Assessor no Gabinete de Juízes do Tribunal Constitucional e Doutorando na Faculdade de Direito da Universidade de Coimbra (joaquim.o.martins@gmail.com / orcid=0009-0006-8485-0898).

RESUMO[*]

O autor analisa, de um ponto de vista prático, a forma como os juízes podem usar a IA nas tarefas que realizam diariamente, procurando determinar em que situações a IA lhes poderá ser útil e aquelas em que deverão evitar recorrer à mesma, concluindo que a IA deverá servir, essencialmente, como uma ferramenta auxiliar para os julgadores, mas nunca para os substituir no âmago e na essência das suas funções jurisdicionais. Numa segunda parte, analisam-se as recentes Recomendações no uso de inteligência artificial do Conselho Superior da Magistratura, considerando que o seu teor é adequado e correto, mas que surgiram com atraso e que será essencial que passem a regras jurídicas vinculativas para juízes, sendo fiscalizada a sua aplicação e sancionado o seu incumprimento, sem que o poderá haver o recurso não autorizado a sistemas de IA pelos juízes, o que poderá colocar em causa a própria legitimidade e legitimação do sistema judicial.

Palavras-chave: Justiça Digital; Inteligência Artificial (IA); Utilização de IA pelos juízes; Recomendações para o uso de IA por juízes; Legitimidade e legitimação do sistema judicial.

---

* Cumpre renovar os agradecimentos ao (agora) Juiz Conselheiro Manuel Ramos Soares, então Presidente da Direção Nacional da Associação Sindical dos Juízes Portugueses, por me ter convidado a coordenar um dos subgrupos (concretamente, o subgrupo relativo à Justiça Digital e à Inteligência Artificial, integrado por Ana Maria Marques Flórido Pinhol, António Joaquim Gomes, Carlos Caiado Pinho, Cláudia Pina de Neves Cunha, João António Ferreira e Manuel David Masseno), do *Think Tank* sobre a Reforma da Justiça, com a coordenação geral do Juiz Conselheiro Nuno Coelho, cujo trabalho depois se corporizou no livro Agenda da Reforma da Justiça (publicado em 2023), mormente no capítulo desse livro relativo a "*Justiça Digital e Inteligência Artificial nos Tribunais*". De facto, até então nada tinha escrito sobre estas temáticas e pouco tinha pensado sobre as mesmas, mas este convite acabou por marcar profundamente o meu percurso profissional e académico nos anos seguintes, como sucede, de novo, com este artigo, em que aproveitei várias das ideias já expressas (e até trechos aí constantes que são reproduzidos quase *verbatim*) na minha entrevista com Adriana Aneli para a série "*O Outro Lado da Tela*", disponível em https://crocodrila.netlify.app/entrevistas/2026-08-jose-joaquim-oliveira-martins/ (consultado, tal como os outros sítios da *Internet*, em 25/8/2026, tendo a tradução de textos sido efetuada, inicialmente, com recurso ao sítio *DeepL* - https://www.deepl.com/pt-PT/translator, mas sempre com revisão final do autor), a quem igualmente agradeço a oportunidade para voltar a refletir sobre esta temática, com perguntas sempre pertinentes e muito desafiadoras, e a forma, sempre simpática e prestável, como foi realizada essa entrevista.

ABSTRACT:

The author analyses, from a practical perspective, how judges can use AI in their day-to-day work, seeking to determine in which situations AI may be useful to them and in which they should avoid relying on it, concluding that AI should essentially serve as an auxiliary tool for judges, but never replace them in the core and essence of their judicial functions. In the second part, the author examines the recent Recommendations on the use of artificial intelligence issued by the Superior Council of the Judiciary, noting that whilst their content is appropriate and correct, they were issued belatedly and that it will be essential for them to become legally binding rules for judges, with their application being monitored and non-compliance sanctioned, as unauthorised use of AI systems by judges must not be permitted, as this could call into question the very legitimacy and legitimation of the judicial system itself.

Keywords: Digital Justice; Artificial Intelligence (AI); The use of AI by judges; Recommendations for the use of AI by judges; Legitimacy and legitimation of the judicial system.

"*About the time Handel became blind, his surgeon, Samuel Sharp (eye surgeon, Guy's Hospital), asked him if he was able to continue playing the organ in public, for the performance of oratorios. Handel replied in the negative. Sharp recommended John Stanley (a blind composer and performer), as a person whose memory never failed; upon which Handel burst into a loud laugh, and said: "Mr. Sharp, have you never read the Scriptures? do you not remember, if the blind lead the blind, they will both fall in the ditch?"* - WILLIAM COXE, *Anecdotes of G.F. Handel and J.C. Smith*, p. 44*

.

I. INTRODUÇÃO

Costuma dizer-se, em jeito de *blague*, que Antonio Vivaldi não compôs seiscentos concertos, mas escreveu o mesmo concerto seiscentas vezes, o que, não sendo minimamente verdade, não deixa de ter alguma razão (*se non è vero, è ben trovato*), dadas as similitudes estilísticas evidentes entre os seus múltiplos concertos, para violino e para tantos outros instrumentos.

De resto, os compositores barrocos eram mestres na reutilização de peças suas ou até de outros compositores, pelo que não se estranha, por exemplo, ouvir uma peça orquestral de Handel (na grafia inglesa) depois reaproveitada para a abertura de uma das suas óperas ou sabermos que Johann Sebastian Bach compôs várias obras a partir, assumidamente, da transcrição de peças de Vivaldi (v.g., Concerto em Ré Menor, BWV 594, com origem no concerto para violino *Grosso Mogul*).

Se o barroco musical se caracteriza também pelo recurso a temas (alguns de origem alheia, como sucede com a *Oferenda Musical* de Bach Pai) e variações sobre os mesmos (como, paradigmaticamente, as *Variações Goldberg* do mesmo compositor), a verdade é que encontramos sempre peças que são objeto de uma espécie de reciclagem *avant la lettre*, podendo servir para abrir uma cantata, depois para andamento intermédio de um concerto grosso e acabar, quiçá, como uma peça para instrumento de tecla. Antecipando os ensinamentos de Lavoisier, químico que nasceu quase no final do período barroco, nada se perde e quase tudo se transforma no mundo da música barroca, havendo, aliás, peças que mantiveram sempre, até hoje, a sua atualidade e fascínio, sendo objeto das mais variadas (e

* Disponível em: https://books.google.pt/books/about/Anecdotes_of_George_Frederick_Handel_and.html?id=0XI_LIVDP9cC&redir_esc=y.

nalguns casos, desvairadas) transcrições ou adaptações, como sucede com as já referenciadas *Variações Goldberg* ou a *Chaconne* de uma das *Sonatas e Partidas para Violino Solo* de Bach, sublimemente transcrita para piano por Ferruccio Busoni.

Aqui chegados, pode parecer (*rectius*, em verdade, é assumida - e até querida - essa estranheza inicial) algo estranho este início de um artigo sobre, pasme-se, IA e juízes, mas a verdade é que tenho passado os últimos tempos a escrever e a falar, nas mais variadas circunstâncias e eventos, sobre as relações entre IA e o sistema judicial, tanto que às vezes parece que me estou já a glosar ou repetir, *auto-plagiando-me*, quase como um demiurgo barroco condenado a repetir sempre o mesmo tema/ária por via de variações e fugas quase sempre similares.

Assim, já algo cansado, resolvi adotar uma diferente abordagem para este artigo – em vez de pegar em tudo o que já escrevi, dar-lhe um novo arranjo formal[1] e aditar algumas novidades (se há matéria em que há sempre desenvolvimentos de última hora, trata-se desta), pensei em partir da minha experiência concreta como juiz, durante mais de vinte anos, e pensar no que, dentro da mesma, poderia ser ajudado (e de que forma) pela IA e, ao mesmo tempo, as matérias em que não se deverá poder recorrer à mesma (repescando também parte das respostas à entrevista referida na primeira nota de rodapé, para permitir que alcancem um público algo diverso), aproveitando igualmente o ensejo para, a final, analisar as recentes recomendações dirigidas aos juízes portugueses da jurisdição comum.

Isto é, fugindo a mais um artigo de matriz académica[2] e em vez de um texto primacialmente teórico e cheio de conceitos mais gerais, ligados à axiologia, à ontologia e à epistemologia (fica, certamente, para outra ocasião…), a ideia será, sem grandes referências doutrinais e citações de textos alheios, dar, de forma mais coloquial (até pela origem de parte do texto aqui revisitado) e algo rapsódica, uma primeira ideia de como a IA pode ser usada no trabalho quotidiano de um juiz na primeira instância (e, pela negativa, como não deve ser utilizada) e

1 Talvez com recurso a IA, o que, frise-se, não aconteceu neste texto – num mundo digital em que somos cada vez mais bombardeados com *slop* e com textos, publicações e comentários tão absurdamente "IAnianos" que parecem, por vezes, paródias não intencionais, é bom, sempre que possível, voltar ao simples labor mental e até artesanal de escrever um texto só por si, sem qualquer assistência digital alheia, que acaba, por muito que possa ajudar, por uniformizar, 'castrar' e 'pasteurizar' muito do que toca, produzindo textos desprovidos de qualquer toque pessoal e humano, frios, superficiais e impessoais, mesmo que elaborados muito rapidamente e aparentemente mais perfeitos a nível formal.

2 Depois de um ano cheio de artigos dessa índole, alguns dos quais foram já publicados ou deverão ser publicados brevemente, desde logo sobre IA e governação de sociedades comerciais, boa fé e dever de lealdade nas sociedades comerciais e propriedade intelectual e jurisdição constitucional.

do que será necessário para evitar passarmos a ter um cenário verdadeiramente distópico a nível dos tribunais digitais que já hoje existem.

II. O TRABALHO QUOTIDIANO DE UM JUIZ E IA

a) Se se pensar um pouco no trabalho diário de um juiz na primeira instância[3], facilmente se alcança que as suas várias tarefas se podem inserir, grosso modo, em três grandes grupos ou tipos: *i)* o presidir a diligências processuais, *ii)* o despachar do expediente processual diário e *iii)* o proferir despachos de fundo/decisões finais.

Assim e de acordo com a sua disponibilidade de agenda (e muitas vezes também da disponibilidade da própria sala de audiências), o juiz preside regularmente a inúmeras diligências inseridas na tramitação de processos, algumas bem rápidas e usualmente simples (como tentativas de conciliação), outras bem complexas e potencialmente demoradas (nos casos mais extremos, com audiências de julgamento com múltiplos meios de prova a serem produzidos e inúmeras sessões, que podem levar um julgamento a durar meses ou até anos).

Quanto ao expediente processual, trata-se normalmente de despachos relativamente simples para assegurar a tramitação e prossecução dos vários processos, i.e., mandando citar partes, agendando, dando sem efeito ou adiando diligências processuais, determinando ou indeferindo a realização de diligências probatórias, atribuindo valores a título de remuneração ou de despesas a intervenientes processuais (peritos, testemunhas), etc.. Este expediente pode referir-se, por sua vez, a processos ainda não decididos (o que será normalmente o caso), mas pode também respeitar a questões que surjam já depois da sua decisão final (como o pagamento de custas, a execução de sentenças, etc.)[4]. Usualmente, tratam-se de despachos aparentemente simples, mas que, pelo seu número (implicando também a consulta da tramitação anterior e dos vários requerimentos juntos desde a última conclusão do processo ao julgador) e pela não simplicidade de algumas das questões a

3 Nas diversas jurisdições que aí existem, dado que, apesar das diferenças patentes que existem entre as mesmas, há também claros pontos de contacto no seu labor quotidiano.

4 Havendo processos em que a litigiosidade entre as partes não se esgota minimamente após a prolação da decisão final, antes se prolonga quase *ad eternum*. desde logo em sede do pagamento de custas e na eventual execução da sentença (recordo-me sempre quando, ainda juiz estagiário, tive de determinar a presença da autoridade policial para a entrega judicial de um terreno, que só ocorreu anos depois de ter sido definitivamente decidido o processo e após várias tentativas infrutíferas para o efeito e de inúmeros incidentes e questiúnculas entretanto suscitados e decididos).

decidir[5], acabam por tirar muito tempo diariamente ao juiz (que, devido a isso e muitas vezes, é obrigado a utilizar o que seria o seu tempo de descanso diário, fins-de-semana e até férias, para a prolação das sentenças mais complicadas, que integram o grupo a que se referirá de seguida).

Relativamente aos despachos de fundo/decisões finais, dizem respeito, como facilmente se alcança, a despachos mais longos que, *verbi gratia*, decidem incidentes da instância (por exemplo, intervenção de terceiros ou determinação do valor da causa), procedimentos cautelares ou incidentes processuais atípicos ou anómalos, a despachos mais complexos (como, paradigmaticamente, despachos saneadores) *ou* a decisões finais (desde logo, sentenças de mérito, com a sua repartição clássica entre fundamentação da matéria de facto e de direito, com o respetivo dispositivo final e condenação em custas).

Assim, grande parte do trabalho quotidiano dos magistrados judiciais não corresponde, ao contrário do que se possa pensar, ao presidir a julgamentos e ao exercício 'puro e duro' da jurisdição, no seu sentido etimológico de 'dizer o direito', antes implica que o julgador (pelo menos em Portugal) realize uma série de tarefas auxiliares ou preparatórias (e que poderiam, na generalidade dos casos, ser assumidas pelos funcionários judiciais ou por assessores), muitas vezes invisíveis e pouco valorizadas.

O dia de trabalho de um juiz não se resume, pois, só a prolatar sentenças numa espécie de vazio e vácuo existencial, antes é intercalado e interrompido por uma inúmera série de tarefas que se diriam menores (e são-no efetivamente), mas sem as quais não seria possível terminar processos e que ocupam imenso tempo ao julgador[6].

b) É evidente também que a IA pode auxiliar muito os juízes nesses dois últimos tipos de tarefas: desde logo, no expediente diário, analisando a tramitação anterior, fazendo um resumo da mesma e propondo um possível despacho (ou até, v.g., datas para diligências de acordo com os anteriores agendamentos e o previsível tempo de duração dessa diligência em

5 Como reclamações a notas de custas de parte ou a consideração do benefício de proteção jurídica, em que se suscitam até várias questões de conformidade constitucional do respetivo regime legal e em que pode haver recurso para tribunais superiores.

6 De resto, pensa-se que seria possível, com algumas alterações legais, aliviar muito desse trabalho secundário judicial, que bem poderia ficar a cargo dos oficiais de justiça, sob direção e orientação do julgador, ou dos assessores judiciais, cuja criação e funcionamento, no seu regime atual, acaba por pouco ajudar a generalidade dos juízes da primeira instância.

concreto), que poderá ser depois assumido (ou não) pelo julgador, que pode, nos casos mais simples, limitar-se a conferir o mesmo e a assiná-lo, assumindo, assim, esse depacho como integralmente "*seu*".

Aliás, o mesmo poderia suceder, *mutatis mutandis*, com os despachos de fundo e as sentenças, em que o sistema de IA podia servir como uma espécie de 'assessor eletrónico', analisando, quanto a estas últimas, os articulados das partes, resumindo os mesmos e concatenando-os no relatório inicial da sentença, referindo quais os factos que foram alegados e que deveriam ser dados como provados ou não provados, fornecendo um esboço de uma possível fundamentação da matéria de facto, pesquisando a legislação aplicável e jurisprudência existente relevante, propondo também uma eventual versão inicial, de acordo com o modelo de decisões anteriores do julgador, da fundamentação da matéria de direito e de possíveis decisões finais.

Em todos estes pontos, a IA serviria sempre como um simples 'assistente eletrónico', facilitando o trabalho do juiz, contribuindo para uma tramitação e decisão mais rápida e eficaz dos processos, cabendo a última palavra sempre ao juiz, pelo que, em regra e desde que cumpridas as condições a que me referirei infra, não me parece que deva ser vedado o recurso a um sistema auxiliar judicial de IA, que não se substitua ao juiz, mas que antes o ajude nessas duas tarefas concretas.

Quanto às diligências processuais e em particular quanto aos julgamentos e adiantando-me já a uma conclusão final (e que ressuma de tudo o que já escrevi a este respeito), parece-me que a situação é totalmente diversa, dado que, desde logo e face a uma série de princípios essenciais do sistema judicial (ainda mais relevantes na própria jurisdição penal), como a imediação ou a plenitude de assistência dos juízes, seria impossível deixar o juiz totalmente "fora" da sala de audiências e deixar um sistema de IA presidir ao julgamento, a gerir a decisão de requerimentos das partes e a apreciar valorativamente a prova produzida[7].

7 Se, na sequência também da pandemia, se assistiram a fenómenos de juízes que vão com pouca frequência aos tribunais onde exercem funções, realizando o seu serviço quase sempre à distância e a partir do seu domicílio (o que não deixa de ter as suas repercussões no quebrar do vínculo entre o julgador, o seu juízo concreto e todos aqueles que trabalham ou recorrem ao mesmo), imagine-se o que sucederia se o julgador se pudesse também substituir na realização de julgamentos – teríamos uma espécie de juiz totalmente virtual, que ninguém vê, mas que trabalha só nas sombras e à distância, o que não se afigura ser minimamente bom para a legitimidade e legitimação do sistema judicial.

Desta forma, afigura-se que, a nível dos julgamentos (e de outras diligências processuais em que haja produção de prova) a intervenção de sistemas de IA pode ser importante para a análise e localização da prova documental já junta (para permitir confrontar as testemunhas com a mesma), para a tradução simultânea (permitindo um contacto mais direto e contínuo com os intervenientes processuais que falem outras línguas que não a portuguesa) e para a transcrição imediata de depoimentos (o que torna possível que, terminado o julgamento, o julgador e todos os sujeitos processuais tenham logo acesso a essa transcrição). Na parte restante, entendo que é indispensável, até pelas disposições constitucionais relativas ao sistema judicial e aos juízes[8], que seja sempre um magistrado judicial a presidir efetivamente a toda a audiência[9], sem recurso a sistemas de IA para lá do que aqueles que, de forma muito limitada, o possam auxiliar nessa tarefa.

c) Quanto à utilização concreta de IA pelos magistrados judiciais, acho imprescindível que o julgador refira sempre expressamente, caso tenha ocorrido, que recorreu à IA e de que forma, *ipso est*, se, por exemplo, foi utilizada para resumir articulados ou para a pesquisa de legislação e jurisprudência, mas assumindo o juiz, em todos os casos, que reviu os respetivos resultados e que se responsabiliza integralmente pelos mesmos, assumindo-os como seus.
Como facilmente se alcança, essa declaração expressa e obrigatória poderá servir para "enfraquecer" a decisão e poderá ser imediatamente utilizada, pelos mandatários das partes, para arguir a sua nulidade em sede de recurso, mas a outra opção - o juiz usar a IA e nada dizer - é claramente pior e poderá até colocar em causa pilares essenciais de qualquer sistema de justiça, incluindo a sua própria legitimidade e legitimação.
O juiz está vinculado, por dever de ofício, a uma série de obrigações e deveres inerentes à sua função - de que se destaca, claro e no que deve ser sempre assegurado em qualquer

8 Que devem administrar a justiça em nome do povo, o que impõe, a meu ver, que não possam, sem mais, ser substituídos, nessa administração da justiça, por sistemas de IA, por mais avançados que sejam, antes competindo sempre aos juízes o efetivo exercício das funções jurisdições que constitucionalmente lhes competem.
9 Diria até mais - sempre me pareceu importante, no meu múnus judicial diário, que os magistrados (judicial e do Ministério Público) e os advogados presentes numa audiência não estejam embrenhados no seu computador e aparentemente sem atenção ao que se diz e ocorre na sua frente, o que é particularmente pernicioso quanto aos juízes. Por esse motivo, foram muito raras as vezes em que levei um computador portátil comigo para a sala de audiências ou em que recorri ao mesmo de forma mais intensiva durante os julgamentos. Nada pior, a meu ver, para uma parte do que ter perante si um juiz que, aparentemente, só está ali de corpo presente e pouco ou nada efetivamente presencia e assiste, dando a ideia até que está ali a realizar uma série de outras tarefas, com pouca (ou nenhuma) atenção à produção de prova.

sistema judicial, a isenção, imparcialidade, etc., mas também a transparência e, como todos nós, a própria honestidade e seriedade intelectual.

O julgador não pode, por exemplo, decidir com base num meio de prova que seja só do seu conhecimento pessoal e a que não se faz referência na própria sentença final, dado que tal violaria as próprias regras processuais e vários dos seus deveres deontológicos. Não me parece que a situação seja diversa quanto à IA - se o juiz usou IA, deve dizer expressamente (e sem medo) que o fez e de que forma, para que os destinatários das decisões percebam de que modo se chegou às mesmas e para que esse uso de IA possa depois ser controlado pelas partes, desde logo em sede de recurso.

Repare-se que a possibilidade oposta, nada dizer, implica sempre que se possa descobrir posteriormente que o julgador omitiu, negligente ou até dolosamente, o ter recorrido à IA, o que me parece ser bem pior a nível da legitimidade e legitimação do sistema judicial e dos juízes que aí exercem funções, como melhor se exporá infra.

De resto e até para salvaguarda da posição do juiz, penso que será melhor que haja essa total transparência de se especificar que se usou IA e de que modo, evitando que erros do sistema de IA sejam assacados, sem mais, a um julgador que se limitou a recorrer à IA e podendo, claro, igualmente servir para acautelar a possibilidade de repetição desses erros em inúmeros processos posteriores e em futuras decisões judiciais[10]. Só sabendo de que forma se atingiu aquele resultado final - *in casu*, a sentença - se poderá compreender e entender a mesma, aferir da sua validade e poder colocar em causa, recursivamente, os seus fundamentos, sem o que se tornará um objeto estranho que "aterra" no final do processo sem se perceber bem porquê e para quê.

Por último, afigura-se que não é o simples recurso a sistemas auxiliares de IA, desde que respeitadas as condições essenciais mencionadas infra, que pode servir para tornar inválida

10 Imagine-se, por exemplo, que o juiz recorreu a um sistema de IA para o cálculo de prestações pecuniárias resultantes de um acidente de trabalho: se é verdade que esse recurso pode poupar muito tempo e trabalho ao juiz, é importante que fique expresso na sentença que se utilizou o mesmo e, mais que isso, que se refiram os dados de base que foram tidos em conta para o efeito (a remuneração anual ou diária, por exemplo) e quais foram, em concreto, as operações matemáticas que conduziram a esse resultado final. Sem essa explicitação, os diversos sujeitos processuais nunca perceberão como se chegou aos valores constantes da sentença e terão uma muito maior dificuldade em questionar os mesmos futuramente, sendo também possível, se se tiver incorrido em algum erro nas operações que levaram aos montantes finais, que o mesmo seja reproduzido em qualquer decisão posterior. Por isso, se acho, até porque estive muitos anos na jurisdição laboral e sei bem o tempo que ocupa esse tipo de cálculos, que é perfeitamente possível que se recorra a IA para esta finalidade, é importante que esse sistema de IA forneça também, para inserção expressa na sentença, os dados de entrada e as operações realizadas e não apenas o resultado final.

ou irregular qualquer decisão em que tenha sido elaborada com o auxílio de IA. Se o julgador assumir o uso de IA e explicitar a forma como o fez, as partes podem sindicar previamente essa utilização e questionar também os seus resultados, impugnando a fundamentação da matéria de facto e os raciocínios probatórios aí constantes e tentando afastar a fundamentação de direito e a respetiva decisão final.

No fundo, trata-se de uma sentença como as outras, apenas tendo sido utilizado um meio tecnológico mais recente auxiliar para a sua elaboração, podendo os sujeitos processuais questionar a licitude do recurso à IA e, mais importante, os seus resultados finais, tudo decorrendo, nesta última parte, como sucede até agora, competindo ao tribunal de recurso aferir da validade da sentença proferida, bem como, agora inovatoriamente, da utilização prévia de IA.

Todavia, é também claro que o uso de IA para lá de meras utilizações auxiliares, que exorbitem das orientações já existente quanto ao seu uso ou que não permitam a sua sindicância em sede de recurso, pode levar à revogação da sentença devido a esse facto. Basta pensar no que sucederá se um julgador, na fundamentação da matéria de facto, referir, tão só, que submeteu os vários meios de prova a um sistema de IA, com os resultados que aí constam expressos, mas sem que se perceba minimamente como se chegou aos mesmos, caso em que a sentença será claramente nula por falta de fundamentação.

d) Finalmente e procurando um critério (muito) prático para saber se um juiz deve (ou não) recorrer à IA, parece-me que será necessário, em primeiro lugar, que o juiz se debruce e analise todas as tarefas que realiza quotidianamente, apurando se a IA o pode ajudar em algumas delas, diminuindo significativamente o seu trabalho e o tempo da realização dessa tarefa *e* assegurando que seja realizada de forma qualitativamente igual ou até melhor do que pelo próprio juiz. Deve-se, assim, evitar recorrer à IA para tudo e para qualquer coisa na tramitação e decisão de um processo, evitando também, *prima facie*, o 'despejar' todo o processo num sistema de IA e pedir que faça todo o trabalho do julgador, aceitando depois acrítica e cegamente o respetivo resultado.

Se se concluir que poderá ser útil o recurso à IA, será depois importante verificar se esse uso da IA, sempre sobre o controlo e responsabilidade final do julgador, respeita os princípios essenciais inerentes a qualquer sistema judicial, incluindo os relativos à utilização da IA, e

os direitos, liberdades e garantias das partes, e não viola as respetivas orientações e regras deontológicas aplicáveis, sendo que só poderá recorrer-se à IA em caso de resposta positiva a estas questões prévias.

Haverá sempre, de todo o modo e na generalidade dos casos, sempre uma área intangível e intocável de reserva de jurisdição do juiz, que não pode (*rectius*, não deve) ser delegada em sistemas de IA, relativa, por exemplo, à produção de prova em sede de julgamento perante o julgador, à valoração e apreciação crítica da prova, à fundamentação da matéria de facto e à própria decisão final, que deverá ser da responsabilidade total do julgador, mesmo que recorra à IA para o auxiliar na sua elaboração.

No fundo, o que se exige ao julgador é que não peça à IA aquilo que a IA não lhe pode dar *e* também que não procure que a IA o substitua nas tarefas que correspondem ao âmago e à essência da função jurisdicional, que deve ser exercida, nessa parte, exclusivamente por juízes-físicos (e nunca robots), mesmo que com algum auxílio da IA, mas cabendo sempre a última palavra ao julgador concreto, que deve assumir integralmente todas as suas responsabilidades jurisdicionais.

e) Quanto à justificação dessa espécie de reserva intangível do julgador, a mesma tem, muito brevemente, uma base ontológica, mas também social e jurídica.

Todos somos, mesmo sem o sabermos, uma espécie de juízes quotidianos. Todos passamos os dias a julgar – os outros, a nós próprios, a factos que aconteceram à nossa frente e a outros de que só ouvimos falar ou sobre os quais lemos.

A nossa vida é feita de inúmeras decisões, mas também de julgamentos implícitos ou expressos e de veredictos pessoais mais ou menos certeiros ou justos. Se a Bíblia avisa que "*Não julguem ninguém e assim Deus não vos julgará!*" (Mateus, 7, 1), é quase impossível que consigamos evitar, no nosso dia-a-dia, formular juízos de valor acerca de tudo o que nos rodeia.

Julgar é, assim, algo que é inerente à própria natureza humana e à nossa vida em sociedade, mas os juízes são pessoas normais que, para além desses 'julgamentos comuns', são encarregues profissionalmente de uma função perfeitamente fora do normal, dado que lhes cabe julgar outras pessoas e os seus litígios e as suas decisões são vinculativas e passíveis de serem executadas coercivamente.

Por seu lado, a IA já entrou de forma decisiva na resolução de pequenos litígios sociais, uma vez que, sem que nos apercebamos, já grande parte dos pequenos litígios de consumo são resolvidos online, sem qualquer intervenção judicial, através das respostas dadas por sistemas de IA a reclamações apresentadas por clientes de, por exemplo, plataformas digitais de prestação de serviços.

Pergunta-se depois é se tal poderá ser replicado em larga escala a nível judicial e a resposta é, como tantas vezes sucede, depende – desde logo, da natureza do processo e dos interesses em causa no mesmo.

Não me repugna que, a nível do direito civil, haja a possibilidade de ser um sistema de IA a decidir pequenos litígios de baixa intensidade e valor em que o réu não se oponha ou até questões cíveis mais simples, como cobranças de dívidas, em que não haja necessidade de produção de prova, podendo depois existir, para evitar erros da IA, reclamação da decisão final para um juiz físico. O mesmo, de resto, pode suceder também a nível da jurisdição administrativa e fiscal, em que existe uma clara litigância de massa, com a repetição de ações com as mesmas questões e que são decididas sem prévia produção de prova, podendo o sistema de IA preparar logo uma minuta de decisão, similar a tantas sentenças anteriores, que o juiz se limitaria a conferir, aceitar (ou não) e tornar sua ao assinar a decisão final proposta. O que já questiono e não aceito é a possibilidade de se replicar esse modelo, de forma cega e acrítica, a nível de outras jurisdições, como, paradigmaticamente, a penal. Desde logo porque, a nível do direito internacional e do direito constitucional, existem outros direitos, liberdades e garantias que só se aplicam quanto ao direito penal, como a presunção de inocência e o princípio in *dubio pro reo*, as várias garantias de defesa e o direito ao recurso, etc., que dificilmente poderiam permitir a decisão exclusivamente por IA desse tipo de processos.

Aliás, mais do que isso, parece-me que há um problema de base e epistemológico quanto à lógica de funcionamento dos sistemas de IA, de base mais probabilística e heurística, e à sua aplicabilidade a processos judiciais, em particular aos penais. Para um juiz, pouco interessará saber se, naquele tipo de processos e crime, a percentagem de condenação aproxima-se dos 100 %, dado que lhe interessa mais a justiça do caso concreto e haverá sempre a possibilidade, por mais alta que seja essa percentagem, de estar perante alguém que não cometeu o crime de que vem acusado, enquanto um sistema de IA dificilmente conseguirá fugir de fazer uma

simples extrapolação estatística do geral para o caso particular. Assim, como fazer com que um sistema de IA saiba o que é a dúvida, quando é antes criado para obter o que se julgam ser certezas absolutas ou irrefutáveis, e como definir a partir de que momento (ou percentagem) essa dúvida pode, por efeito do princípio já referido, levar à absolvição (ou não) de um acusado?

Finalmente, há também algo a que não podemos escapar, dado que resulta da própria *natura rerum*.

Os juízes são homens e mulheres como todos os outros e como tal, parafraseando Terêncio, nada do que é humano lhes pode alguma vez ser estranho, pelo que podem, desde logo, sentir empatia e colocar-se até no lugar de quem é julgado ou teve intervenção nos factos submetidos a juízo, num exercício de alteridade, o que lhes dá uma visão e uma mundividência bem diversa de um sistema de IA.

Julgar nunca é, para um juiz, algo de asséptico, frio ou desprovido de consequências, não se podendo confundir isenção e imparcialidade com falta de empatia e humanidade por quem é julgado ou por quem está a ser ouvido num julgamento.

O juiz procura ser o mais objetivo possível, mas não esquece nunca que está a julgar pessoas e que deve ser o mais justo possível para as mesmas dentro das regras legais, mas atendendo igualmente ao sentido de equilíbrio, ponderação e bom senso que todo o (bom) julgador deve ter. As soluções legais mais evidentes podem, por vezes, ser profundamente injustas e inadequadas e é para isso que serve, por exemplo, a declaração de inconstitucionalidade de uma lei ou a recurso a institutos jurídicos que são uma espécie de válvula de escape do sistema jurídico, como o abuso de direito, a que dificilmente um sistema de IA iria recorrer uma vez que se limitaria, no fundo, a replicar automaticamente decisões anteriores.

O juiz não é apenas um técnico do direito, mas é antes também um 'homem/mulher-bom' que, mais do que realizar silogismos legais numa espécie de vácuo existencial axiologicamente neutro, procura que se faça justiça efetiva e que os destinatários dessa justiça a compreendam e sintam que foi feita, enquanto um sistema de IA sempre será alheio a esse tipo de preocupações e finalidades mais profundas do sistema judicial, limitando-se a reproduzir o anteriormente decidido e 'fossilizando' o próprio direito e a jurisprudência. Ser juiz é também, muitas vezes, mudar de opinião, revertendo posições anteriores e explicando

porquê, bem como também decidir contra a jurisprudência fixada, que acaba, muitas vezes e passados alguns anos, por ser alterada em sentido oposto[11].

Por último, é também claro que o próprio simbolismo dos julgamentos e a interação entre os sujeitos processuais se perderia totalmente se o julgador for um sistema de IA. Os julgamentos podem ser vistos como uma interação social performativa e até teatros (e são-no muitas vezes), mas cumprem uma série de funções simbólicas e sociais e os juízes estão ali justamente por serem pessoas, por poderem ouvir presencialmente as partes e por representarem o povo, decidindo a final e dizendo de viva-voz, perante as partes e quem assista a esse ato público, o porquê das suas decisões, contribuindo para encerrar e pacificar o litígio social subjacente. Neste ponto, tal como noutras funções - como os sacerdotes, os médicos ou os psicólogos - o facto de o juiz ser uma pessoa física é indispensável para que haja esse contacto e interação humanos, permitindo, por exemplo, que uma parte desabafe tudo o que lhe vai na alma ou que um réu possa defender-se até ao limite das suas forças, sempre perante alguém como eles, de carne e osso e que também sofre e ri, e que está ali exclusivamente para os ouvir e depois decidir.

Em resumo, julgar é um ato profundamente humano e até ontológico, que assenta no que somos enquanto espécie e enquanto seres sociais, cumprindo igualmente, nos tribunais, uma função social muito necessária e que só pode, em regra, ser realizada por pessoas físicas, pelo que o juiz dificilmente poderá ser substituído, em todos os processos, por sistemas de IA, ganhando todos nós, pelo contrário, se se mantiver esse "toque humano" na maior parte dos casos, em particular nos processos de natureza criminal mais grave ou nos processos em que estejam em apreciação interesses mais relevantes ou valiosos (como processos de família e menores).

Em suma, a digitalização judicial veio para ficar e a mesma já me ajudou muito nas minhas tarefas enquanto juiz, sendo utópico pensar que vamos voltar totalmente ao tempo dos velhos 'tribunais analógicos', mas sou juiz e de uma geração que atravessou parte da transição digital, com todas as dificuldades que implicou, pelo que estou bem ciente de que haverá

11 Só para exemplo, vejam-se as alterações jurisprudenciais relativos ao conceito de 'terceiros para efeito de registo', que levou à fixação sucessiva de jurisprudência em sentidos diversos pelo Supremo Tribunal de Justiça e o resumo constante de RITA SOFIA DUARTE NOBRE, *Resenha jurisprudencial sobre o conceito de "terceiros" para efeitos de registo*, https://repositorio.ucp.pt/bitstreams/c7d3b05f-d317-4794-aaf2-1b5673120d47/download.

sempre inúmeros riscos com a IA judicial, desde logo, como se exporá melhor infra, o já famoso *deskilling*, o efeito de ancoragem ou, num cenário mais distópico, a própria conversão dos juízes numa espécie de 'relações-públicas' de sistemas de IA judicial, anunciando só os seus resultados, sem verdadeiramente os determinarem ou sequer influenciarem.

Esses riscos, bem patentes, só podem ser evitados pela vinculação (e assunção) dos juízes a uma série de deveres deontológicos, passando, de seguida, a analisar um dos instrumentos normativos (ou talvez só 'recomendatório') mais recente e relevante para os julgadores portugueses.

III. ÉTICA, DEONTOLOGIA JUDICIAL E IA - O CASO PORTUGUÊS

a) Face à importância e relevância da IA no próprio sistema judicial, abundam, como se sabe, os instrumentos de *hard* e, em especial, *soft law* relativos à utilização de IA pelos juízes, como, mais recentemente, as *Recomendações no uso de inteligência artificial do Conselho Superior da Magistratura* (CSM)[12].

Estas *Recomendações*, mesmo que muito tardias, assentam em considerandos gerais que são perfeitamente acertados, embora deva-se referir, quando aí se escreve que "*iv) Nos termos do art.23.º da Lei n.º 34/2009, de 14 de julho, a responsabilidade pelo tratamento de dados nos processos judiciais, cabe ao CSM e aos juízes com competência sobre o respetivo processo*", que "*é necessário efetuar uma (sempre difícil) concordância prática entre as finalidades e os pressupostos essenciais de qualquer sistema judicial, que dizem respeito, por exemplo, aos direitos fundamentais garantidos aos sujeitos processuais (como, desde logo, o direito ao contraditório), e a tutela dos dados pessoais que estão recolhidos nos respetivos processos judiciais. Seria, assim, necessário que em Portugal se previsse a forma de execução no sistema judicial do regime constante no RGPD, o que ainda não sucedeu e causa bastantes dificuldades quanto ao modo concreto como esse Regulamento se aplica*

12 https://csm.org.pt/wp-content/uploads/2026/05/Recomendacoes_Uso_IA.pdf. Estas recomendações foram aprovadas pelo Plenário do CSM em 8 de abril de 2026, mais de um ano após a aprovação do *Projeto da Carta Ética para a Inteligência Artificial nos Tribunais Administrativos e Fiscais* (Anúncio (extrato) n.º 95/2025, publicado no Diário da República n.º 66/2025, Série II de 3/4/2025) - https://cstaf.pt/wp-content/uploads/2025/03/T006_Carta_Etica_Draft-8.pdf, e, curiosamente, justamente no mesmo dia em que foi aprovada, pelo Conselho Superior do Ministério Público, a *Carta para a utilização ética da IA pelo Ministério Público e Política de Auditoria e Monitorização Técnica* – https://www.ministeriopublico.pt/sites/default/files/editor-files/ia_carta_etica-politica_auditoria.pdf.

nesta matéria", sendo que, citando MANUEL DAVID MASSENO, "*lamentavelmente, é ainda a Lei n.º 34/2009 de 14 de julho, que "Estabelece o regime jurídico aplicável ao tratamento de dados referentes ao sistema judicial", o que "coloca problemas especialmente complexos e controversos, sobretudo no respeitante à determinação do(s) responsável(eis) pelo tratamento do dados, com a(s) correspondente(s) responsabilidade(s) proativa(s) (accountability, Art.º 5.º n.º 2, 24.º, 25.º e 32.º do RGPD) e a inerente imputabilidade das responsabilidades civil (Art.º 82.º do RGPD e Art.º 33.º da Lei n.º 58/2019, de 8 de agosto, assegura a execução, na ordem jurídica nacional, do Regulamento (UE) 2016/679) e contraordenacional (Art.º 83.º do RGPD, Art.ºs 37.º a 45.º da Lei n.º 58/2019 e, ainda, a Lei n.º 67/2007, de 31 de dezembro, aprova o regime da responsabilidade civil extracontratual do estado e demais entidades públicas)*"[13].

No mais, o CSM deliberou estabelecer um "conjunto de recomendações quanto à utilização dos sistemas de IA por juízes dos tribunais judiciais no exercício da atividade jurisdicional" (Art. 1.º), aí constando a, final, um glossário (útil) de definições, devendo essas recomendações e esse glossário serem revistos regularmente, "Considerando a evolução tecnológica e de experiência de utilização", e sendo essas recomendações divulgadas pelos juízes da jurisdição comum (Art. 2.º, respetivamente, n.os 1, 2 e 3).

Quanto aos princípios aí constantes, no seu Art. 3.º, são os seguintes: "*1. Princípio do controlo humano eficaz. A utilização de sistemas de IA no exercício da atividade jurisdicional estará sempre sujeita ao controlo humano real, consciente e eficaz por parte dos juízes, não podendo tais sistemas operar autonomamente para a tomada de decisões judiciais, a avaliação de factos ou provas, ou a interpretação e aplicação da lei. 2. Princípio da não substituição dos juízes. No exercício da atividade jurisdicional, a utilização de sistemas de IA não poderá, em caso algum, substituir os juízes na tomada de decisões judiciais, na avaliação de factos ou provas, ou na interpretação e aplicação da lei. 3. Princípio da responsabilidade judicial. A responsabilidade plena e exclusiva pelos*

13 JOSÉ JOAQUIM FERNANDES OLIVEIRA MARTINS, «Proteção de dados pessoais, privacidade e sistema Judicial português» in *Revista do Centro de Estudos Judiciários*, 2.º semestre de 2003, II, pp. 226-227, (v. também, já antes, *Proteção de Dados e o Sistema Judicial Português – Uma síntese* in AA.VV., organização de LUCIANE CARDOSO BARZOTTO/RICARDO HOFMEISTER DE ALMEIDA MARTINS COSTA, *Estudos sobre LGPD – Lei Geral de Proteção de Dados – Lei n.º 13.709/2018: Doutrina e Aplicabilidade no Âmbito Laboral*, Escola Judicial do Tribunal Regional do Trabalho da 4.ª Região, Porto Alegre, 2022, disponível em https://lume.ufrgs.br/handle/10183/238583), sendo que deverá ter em conta, mais recentemente, o "Acordo de Tratamento de Dados Pessoais" a que se aludirá infra.

despachos, sentenças e acórdãos judiciais recai, em todos os casos, sobre os juízes, independentemente da utilização dos sistemas de IA como ferramenta de apoio ou assistência. 4. Princípio da independência judicial. A utilização de sistemas de IA deve ser realizada de forma compatível com a independência judicial, sem que os resultados gerados por estes sistemas condicionem, direta ou indiretamente, a liberdade de julgamento do órgão judicial. 5. Princípio do respeito pelos direitos fundamentais. A utilização de sistemas de IA deve, em todos os casos, respeitar os direitos fundamentais reconhecidos na Constituição e no direito da União Europeia e, em particular, o direito à tutela jurisdicional efetiva, à igualdade, à não discriminação e à proteção dos dados pessoais. 6. Princípios da confidencialidade e da segurança. No tratamento de dados judiciais com recurso a sistemas de IA, a confidencialidade, a integridade e a segurança da informação devem ser garantidas, prevenindo o acesso não autorizado, a utilização indevida ou as transferências não autorizadas. 7. Princípio da prevenção dos vieses algorítmicos. Os juízes devem tomar as precauções necessárias para identificar e evitar os enviesamentos algorítmicos decorrentes da utilização de sistemas de IA. 8. Princípios da proporcionalidade e da utilização limitada. A utilização de sistemas de IA deve ser proporcional à finalidade pretendida e limitada aos casos em que possam ser úteis e eficazes como ferramenta de apoio ou auxílio à atividade jurisdicional. 9. Princípio da formação e capacitação. Os juízes devem receber formação e capacitação sobre a utilização de sistemas de IA na administração da justiça. 10. Princípio da transparência. O uso de sistemas de IA deve ser assumido e declarado nas decisões judiciais".

Relativamente a estes princípios, claro que *nihil sub sole novum*, vindo, manifestamente (e até expressa e assumidamente), na sequência da *Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente adotada pela CEPEJ (Comissão Europeia para a Eficácia da Justiça)* na sua 31.ª reunião plenária (Estrasburgo, 3 e 4 de dezembro de 2018)[14] e do próprio *Código Ético Ibero-Americano para o Uso de Inteligência Artificial no Judiciário*[15], com pequenas alterações de designação e formulação, sem que haja mais propriamente (e ainda bem…) grandes novidades a esse respeito.

14 https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0.

15 Constante, na sua tradução portuguesa, em https://www.cumbrejudicial.org/sites/default/files/2025-05/C%C3%B3digo%20%C3%89tico%20Ibero-Americano%20para%20o%20Uso%20de%20IA%20%28PT%29.pdf.

O CSM recomenda, por seu lado, que "os juízes apenas utilizem os sistemas de IA, incluindo as ferramentas de IA generativa, que lhes sejam disponibilizados pelo sistema de tramitação processual previsto na lei processual ou pelo Conselho Superior da Magistratura", que "serão sujeitos a controlo de qualidade e auditoria pelo Conselho Superior da Magistratura (Art. 4.º, n.os 1 e 2). Esses sistemas podem ser utilizados apenas "como ferramentas de apoio ou de assistência", dado que "Em particular, os juízes podem utilizar os sistemas de IA para os seguintes fins: a) Pesquisa e localização de informação jurídica, incluindo a identificação de normas, jurisprudência e doutrina jurídica relevantes. b) Análise, classificação e estruturação de informações, documentos ou dados constantes de processos judiciais, para efeitos de organização, compreensão ou apoio ao estudo do caso. c) Elaboração de rascunhos, sumários ou minutas internas de trabalho, desde que não sejam decisivos e não substituam a redação pessoal das decisões judiciais. d) Apoio a tarefas organizacionais ou auxiliares relacionadas com a gestão do conhecimento jurídico ou a preparação de trabalhos judiciais" (Art. 5.º, n.os 2 e 3).

Por seu lado, nos termos do Art. 6.º, "*1. Os juízes só devem utilizar minutas geradas por sistemas de IA que lhes sejam fornecidas pelo sistema de tramitação processual previsto na lei processual ou pelo Conselho Superior da Magistratura como ferramentas de apoio ou assistência ao exercício da sua atividade jurisdicional, em plena conformidade com estas recomendações. 2. No exercício da atividade judicial, a utilização de minutas de decisões judiciais ou processuais geradas pelos sistemas de IA exigirá, em todos os casos, a conformidade com os critérios, as orientações de utilização e os princípios estabelecidos nas presentes recomendações. 3. Independentemente da utilização, será necessária uma revisão e validação pessoal, completa e crítica por parte do juiz, que manterá sempre a sua responsabilidade exclusiva pela decisão correspondente. 4. Os projetos gerados por sistemas de IA para o exercício da atividade judicial não devem ser utilizados sem revisão final. 5. Para os efeitos previstos nos números anteriores, os sistemas de IA utilizados no exercício das funções judiciais devem assegurar que o documento preliminar é gerado apenas por vontade do juiz e pode ser livre e integralmente modificado por este antes da sua validação como decisão judicial*".

Após os usos "recomendados" seguem-se, delimitando os mesmos agora pela negativa e no seu Art. 7.º, os "Usos não recomendados" - "*No exercício da sua atividade judicial, os juízes*

não devem utilizar os sistemas de IA previstos na recomendação quatro para os seguintes fins: a) A substituição, automatização ou delegação da tomada de decisões judiciais, da avaliação de factos ou provas, ou da interpretação e aplicação da lei. b) A utilização de resultados gerados por sistemas de IA que, direta ou indiretamente, condicionem a independência judicial ou a liberdade de julgamento dos juízes. c) A incorporação em decisões judiciais de conteúdos gerados por sistemas de IA sem validação crítica, completa e pessoal por parte dos juízes. d) A utilização de sistemas de IA para o tratamento de dados pessoais especialmente protegidos ou de informações sujeitas a obrigações de confidencialidade reforçadas, fora dos casos expressamente autorizados pela legislação em vigor. e) A utilização de sistemas de IA para a elaboração de perfis de indivíduos, previsão de comportamentos, avaliação de riscos ou classificação de sujeitos, fora dos casos expressamente autorizados pela legislação em vigor. f) Quaisquer outras utilizações contrárias aos critérios, orientações de utilização e princípios estabelecidos nas presentes recomendações", bem como, agora no Art. 8.º, os "Sistemas de IA não recomendados", dado que "*No exercício da atividade jurisdicional, os juízes não devem utilizar sistemas de IA, incluindo ferramentas de IA generativa, a menos que lhes tenham sido fornecidos pelo sistema de tramitação processual previsto na lei processual, nas condições previstas no n.º 2 da recomendação 5.º das presentes recomendações, ou pelo Conselho Superior da Magistratura*", sendo que "*Não obstante o disposto no número anterior, os juízes podem utilizar estes sistemas de IA para fins de preparação ou estudo, como a elaboração de resumos ou traduções, ou a análise de fontes jurídicas, doutrinais ou técnicas, desde que a informação utilizada para o efeito provenha única e exclusivamente de fontes abertas*"[16] (n.os 1 e 2, dispondo os números seguintes que "*3. Os juízes não podem, em caso algum, fornecer ou utilizar dados judiciais nestes sistemas de IA. 4. Na utilização destes sistemas de IA o uso de qualquer elemento processual deve estar devidamente anonimizado.*"

O Art. 9.º é relativo à proteção de dados pessoais, remetendo, desde logo, para a "legislação aplicável" (n.º 1) e para princípios e regras já resultantes do Regulamento Geral sobre a Proteção de Dados (n.os 2 e 4), sendo que "*Para efeitos das presentes recomendações, os dados pessoais a que os juízes têm acesso no exercício da sua atividade judicial não podem*

16 Com algumas cautelas: "*Os juízes não podem, em caso algum, fornecer ou utilizar dados judiciais nestes sistemas de IA*" e "*Na utilização destes sistemas de IA o uso de qualquer elemento processual deve estar devidamente anonimizado*".

ser utilizados em sistemas de IA que não tenham sido disponibilizados pelo sistema de suporte à atividade dos tribunais, mediante prévia validação pelo CSM ou por este diretamente fornecidos" (n.º 3) e que "*Para os efeitos previstos nas recomendações anteriores, os sistemas de IA utilizados pelos juízes no exercício da sua atividade jurisdicional devem incorporar medidas técnicas e organizativas adequadas que garantam a segurança dos dados tratados, a prevenção do acesso não autorizado, a rastreabilidade das operações realizadas e a impossibilidade de reutilização dos dados para fins diferentes daqueles para os quais foram tratados*" - n.º 5 (embora, agora nos termos do n.º 6, "*Presume-se que os sistemas de IA fornecidos ou validados pelo CSM têm as características referidas no número anterior*").

O Art. 10.º impõe transparência no uso de IA, dado que, como já se defendeu supra, "*A utilização de sistemas de IA na atividade jurisdicional deve ser acompanhada de referência expressa à sua utilização no despacho ou sentença*" (n.º 1[17]), prescrevendo o Art. 11.º que "*Para os efeitos previstos nas presentes recomendações, caberá ao Conselho Superior da Magistratura exercer as funções de supervisão e controlo sobre a utilização dos sistemas de IA em todas as questões que afetem o tratamento de dados pessoais efetuado para fins jurisdicionais*", enquanto que o Art. 12.º prevê um dever especial de vigilância dos juízes, "*Reconhecendo o carácter de elevado risco do uso de IA na atividade jurisdicional devem os juízes, enquanto responsáveis pelos dados judiciais do processo, prestar especial atenção ao uso de IA, designadamente nas seguintes situações: a) informações jurídicas obtidas por intervenientes sem representação por profissional forense; b) produção de peças processuais com indícios de uso de IA*" (n.º 1)[18].

Last, but not least, o Art. 13.º, sob a epígrafe "Formação e Capacitação", dispõe, no que será mais uma norma programática e autovinculadora, que "*O CSM fornecerá aos juízes atividades de formação e treino sobre a utilização de sistemas de IA*".

17 Afigurando-se, indo até mais longe do que aí se prevê, que o juiz não se limite só a declarar que recorreu a IA de acordo com estas *Recomendações*, mas também de que forma específica ocorreu esse recurso à IA e relativamente, por exemplo, a que partes concretas de uma sentença (na elaboração do relatório, na análise da prova documental, etc.).

18 Sendo que "*No exercício deste dever podem ser considerados indícios do uso de IA, designadamente: a) Citação de acórdãos sem existência real e produzidos por alucinação do algoritmo; b) Citação de legislação sem aplicação em território nacional". c) O uso de expressões não correntes em português europeu; d) Indícios de respostas do algoritmo a prompts*" - Art. 12.º, n.º 2.

b) Cumpre, pois, efetuar uma breve análise crítica destas recomendações, devendo, desde logo, destacar-se que, em geral, são perfeitamente corretas e adequadas as orientações aqui fixadas, que vêm, aliás, na sequência exata do muito que tem sido escrito por vários autores a este respeito (nos quais me incluo, *prima facie* no respetivo capítulo do livro da *Agenda da Reforma da Justiça* e em múltiplos artigos posteriores), não se deixando, todavia, de lamentar que se tenha aguardado até 2026 para entrarem em vigor estas recomendações.

Por sua vez, não se pode deixar de mencionar o (enorme) elefante no canto da sala: o facto de não serem ainda os Conselhos Superiores a terem a total gestão (ou até sequer a cogestão) do sistema informático dos tribunais[19].

Como escrevi anteriormente: "*a titularidade e gestão das redes de dados, sistemas e programas informáticos em que é efetuada a tramitação eletrónica dos processos judiciais compete ao Ministério da Justiça, por intermédio do Instituto de Gestão Financeira e Equipamentos da Justiça, I.P., que os tem integralmente na sua tutela e controlo. Ora, não se vê como podem os juízes ser totalmente independentes no seu exercício de funções quando todos os instrumentos informáticos que usam para o efeito são geridos pelo poder executivo, que, por exemplo, escolhe os equipamentos informáticos a atribuir aos próprios juízes e desenvolve, com muito pouco acolhimento das sugestões dos juízes, os programas informáticos que são usados para tramitar processos. Esta dependência total dos juízes pode comprometer a sua independência, dado que não têm total controlo (em verdade, não têm quase qualquer controlo) sobre a forma como trabalham digitalmente e as modificações*

19 Apesar de, como escreve MANUEL DAVID MASSENO, *Da segurança no tratamento de dados pessoais pelo Poder Judicial em Portugal e no Brasil,* p. 122-123, https://www.iuris.edu.pt/xms/files/Cyberlaw_by_CIJIC_ed_XIII_versao_final.pdf,"*Entretanto, face ao impasse legislativo, foi intentada uma "via criativa", sobretudo destinada a legitimar a anterior situação de facto. Sendo certo que a via mais consentânea com a Constituição e o RGPD teria sido a da transferência das competências e dos meio materiais relativos aos sistemas e às redes informáticas da Justiça para o CSM e para o CSTAF. Assim, no dia 5 de dezembro de 2023, foi assinado um "Acordo de Tratamento de Dados Pessoais", tendo por objeto tanto a Justiça Cível quanto a Criminal, pelos Presidentes do CSM e do IGFEJ, aliás coadjuvados pelas correspondentes EPDs. […] Conforme ao Acordo, o IGFEJ passou a assumir formalmente a posição de "subcontratante" [processador], tratando os dados por conta do CSM, enquanto "responsável pelo tratamento" [controlador] (Artigo 4.º 7) e 8) e 28.º n.ºs 1 e 3 do RGPD). Em especial, o primeiro assumiu obrigações quanto à aplicação de medidas técnicas e organizativas destinadas a efetivar a segurança dos dados objeto de tratamento, assim como no respeitante à notificação das violações de dados pessoais ao CSM, assistindo ainda este na comunicação de tais violações aos titulares dos dados (Cláusulas Sexta, Décima Primeira e Décima Segunda e Décima Terceira)*". Embora este acordo seja relevante quanto à proteção de dados pessoais, o problema de base e legal mantém-se perfeitamente inalterado, desde logo quanto ao controlo e gestão do próprio sistema informático dos tribunais, que se mantém totalmente fora da esfera de competência do CSM.

ao respetivo sistema informático, que podem até colocar em causa o seu cabal e normal exercício de funções"[20].

Sem essa transferência de responsabilidades, estando na disponibilidade de uma entidade totalmente alheia aos Conselhos Superiores a gestão dos respetivos sistemas informáticos dos tribunais, dificilmente poderão esses Conselhos, por exemplo, assegurar que os dados utilizados em sistemas de IA judicial sejam totalmente fidedignos ou garantir, por si só e como consta destas recomendações, a "*segurança dos dados tratados, a prevenção do acesso não autorizado, a rastreabilidade das operações realizadas e a impossibilidade de reutilização dos dados para fins diferentes daqueles para os quais foram tratados*".

Efetivamente, com estas recomendações começa-se, perdoe-se a metáfora, a casa pelo telhado, dado que o substrato fundacional mais básico não existe, continuando a ser o poder executivo a ter a tutela e controlo exclusivo do sistema informático judicial, sobre o qual depois poderão correr aplicações de IA fornecidas pelo CSM, que se desconhece ainda quais sejam e de que forma se vão articular e integrar (se é que tal vai ser possível) com a infraestrutura informática de base sobre a qual o CSM não tem qualquer controlo.

Por outro lado, a maior debilidade destas "*Recomendações*" é justamente o facto de o serem*, ipso est*, de não terem, aparentemente, qualquer vinculatividade ou, pelo menos, a mesma não ser assumida *qua tale*, não estando também previstas quaisquer consequências para o seu incumprimento.

Ficam, assim, uma série de questões a esse respeito – se se dão uma série de recomendações aos juízes e estes não as cumprirem, pode haver consequências para os magistrados judiciais dos tribunais comuns? E, se sim, em que sede, a nível disciplinar ou classificatório? Ou são apenas um "*documento recomendatório*", uma espécie de apelo e exortação bem intencionados, mas não mais do que isso, para apelar à consciência ética dos juízes, sem qualquer eficácia e efeito jurídico?

Assim, por referência à *Carta Ética para a Inteligência Artificial nos Tribunais Administrativos e Fiscais*, mas com aplicação, apesar de algumas melhorias aí constantes, a estas *Recomendações* bem posteriores, entende-se que "*Apesar dos aspetos positivos, a Carta Ética evidencia diversas limitações. A sua natureza meramente recomendatória, sem*

20 JOSÉ JOAQUIM FERNANDES OLIVEIRA MARTINS, «Justiça Digital e Inteligência Artificial - Um Admirável Mundo Novo?» in *Julgar* n.º 52, janeiro-abril de 2024 - "O 25 de Abril e a Justiça".

força legal vinculativa, compromete a sua efetividade e aplicabilidade prática. A linguagem adotada, embora ética e intencionalmente inclusiva, revela-se por vezes demasiado abstrata, dificultando a operacionalização dos princípios nela contidos. A ausência de mecanismos claros de fiscalização e de responsabilização para violações das diretrizes éticas torna o documento vulnerável à ineficácia. Além disso, a Carta Ética não apresenta um modelo de avaliação de riscos específico para as tecnologias de IA nem critérios objetivos para aferir os impactos sobre os direitos fundamentais. Por fim, verifica-se uma débil articulação com os instrumentos normativos europeus e internacionais em vigor, o que limita o seu alcance no contexto de um espaço jurídico cada vez mais globalizado" [21].

Seguindo com estes autores e novamente com aplicação *in casu*, "*Para que a Carta Ética possa ser efetivamente implementada e gerar confiança na sociedade, propõe-se o desenvolvimento das seguintes hipóteses: (i) a sua elevação a instrumento vinculativo com base legal; (ii) a criação de um organismo de supervisão ética com poder de fiscalização e sanção; (iii) a definição de um sistema de classificação de riscos para IA no setor da Justiça; (iv) a inclusão de obrigações de formação contínua para magistrados e funcionários judiciais; (v) o alinhamento explícito com o RIA e com os tratados internacionais em vigor; e (vi) a institucionalização de mecanismos de participação cidadã no desenho, avaliação e controlo da IA na área da Justiça*"[22].

21 RICARDO LOPES DINIS PEDRO/PEDRO MIGUEL ALVES RIBEIRO CORREIA/SUSANA ANTAS VIDEIRA, *Inovações, inteligência artificial e tecnologias de informação e comunicação em Sistemas de Justiça*, p. 12, https://enajus.org.br/anais/assets/papers/2025/sessao-9/a-regulacao-do-uso-de-inteligencia-artificial-nos-tribunais-administrativos-e-fiscais-portugueses-quo-vadis.pdf.

22 RICARDO LOPES DINIS PEDRO/PEDRO MIGUEL ALVES RIBEIRO CORREIA/SUSANA ANTAS VIDEIRA, *ob. cit.*, p. 14.

IV. CONCLUSÃO

Como se sabe e mais recentemente, tem existido um grande enfoque sobre os deveres éticos e de conduta dos juízes, desde logo com o recente Código de Conduta dos Juízes dos Tribunais Judiciais, que é, na verdade, uma espécie de mini(malista)-Código de Conduta, mais 'para inglês ver'[23] (*rectius*, em verdade, para o GRECO - *Group of States against Corruption* - ver[24]). De facto, para além de conceitos demasiadamente gerais ou de muito difícil concretização[25], mistura-se nesse Código o que são condutas claramente com

23 Até pelo facto de só ter havido, desde então, um único Parecer do Conselho de Ética (criado em resultado desse Código de Conduta) em resposta a um pedido de um juiz - https://csm.org.pt/wp-content/uploads/2025/05/conselho-etica-parecer_1-2025.pdf.

24 V. o Segundo Relatório Intercalar de Conformidade Portugal adotado pelo GRECO na sua 87.ª Reunião Plenária (22-25 março 2021), p. 13, https://rm.coe.int/quarto-ciclo-de-avaliacao-prevencao-da-corrupcao-em-relacao-a-deputado/1680a21607, com itálico adicionado - O GRECO toma nota da informação facultada. Não pode concordar que as disposições do Estatuto dos Magistrados Judiciais revisto (que o GRECO pôde consultar) correspondam a um código de conduta claro e aplicável, abrangendo questões como ofertas e conflitos de interesses, tal como é exigido pela recomendação. Parece também que, por enquanto, existem vários outros textos sobre ética no seio do poder judicial, mas a sua relevância e conformidade uns com os outros parece pouco clara. O GRECO conclui que esta parte da recomendação ainda não foi implementada. "*A este respeito, recorda igualmente que o artigo 19.º, n.º 3 da Lei n.º 52/2019 obriga ambos os conselhos superiores a adotar, tendo em devida conta a sua autonomia, códigos de conduta, de acesso público, para os juízes, abrangendo, entre outros, as ofertas e a hospitalidade. O GRECO insta as autoridades a acelerar o cumprimento da primeira parte da recomendação, uma vez que esta é indispensável para o cumprimento da segunda parte da recomendação."* Além disso, em relação ao aconselhamento ao dispor dos juízes, o GRECO não recebeu a garantia de que existam regras que impeçam os membros dos conselhos superiores que prestam aconselhamento sobre dilemas éticos aos juízes de participarem em eventuais processos disciplinares relativos aos mesmos juízes".

25 Só um pequeno exemplo - no Art. 3.º escreve-se, copiando acrítica e expressamente um conceito constante de um dos comentários do *Compromisso Ético dos Juízes Portugueses* (p. 14-15, https://asjp.pt/assets/files/imported/compromisso_etico_dos_juizes_portugueses-5466c7fabd.pdf, - "Em especial, o juiz abstém-se de aderir a organizações coletivas e de participar em debates públicos, sempre que, segundo a apreciação de uma *pessoa razoável, bem informada, objetiva e de boa fé*, isso possa perturbar a imagem de imparcialidade ou independência relativamente a questões suscetíveis de vir a ser submetidas aos tribunais") que "Os juízes dos Tribunais Judiciais abstêm-se de participar em atividades extrajudiciais que possam ser considerados, por uma *pessoa razoável, bem informada, objetiva e de boa-fé*, como suscetíveis de afetar a confiança dos cidadãos na imparcialidade das suas análises e decisões" (itálicos adicionados), fazendo referência uma espécie de "Super Bom Pai de Família", 'razoável, bem informado, objetivo e de boa-fé' - o que quer que isso seja…, de muito difícil concretização e aferição. Essa expressão foi depois recuperada, curiosamente, no Código de Conduta para os membros do Conselho Superior da Magistratura, em que, frise-se e tal como sucede neste Código para todos os juízes, nada se refere quanto às temáticas, igualmente importantes quanto àquelas que aí constam – ou até mais –, do assédio, em todas as suas vertentes, especialmente relevantes para quem desempenha funções deste tipo. Como já escrevi, em conjunto como uma colega, anteriormente, "face à alteração operada pela Lei n.º 73/2017, de 16 de agosto, no Artigo 71º, n.º 1 da Lei Geral do Trabalho em Funções Públicas (aplicável subsidiariamente aos magistrados judiciais por força do disposto no Artigo 188º do atual Estatuto dos Magistrados Judiciais), os Conselhos Superiores (que são, na judicatura, o que mais aproximado os juízes têm aos empregadores públicos aí referidos) devem: "(…) k) Adotar códigos de boa conduta para a prevenção e combate ao assédio no trabalho e instaurar procedimento disciplinar sempre que tiver conhecimento de alegadas situações de assédio no trabalho" (ANA MICAELA SOARES MARQUES

relevância disciplinar[26], correspondendo à violação de deveres essenciais para qualquer juiz, com matérias éticas, sendo o único artigo que é verdadeiramente relevante o artigo 5.º, relativo a "Ofertas, convites e hospitalidade"[27] (até por corresponder às exigências do GRECO).

Ora, seria importante que estas recomendações fossem, em parte, recolhidas nesse Código de Conduta (a nível da transparência do uso da IA, por exemplo) e, mais do que isso, que, no que diz respeito às condutas mais graves, fossem mesmo vinculativas e de cumprimento obrigatório para os juízes, com efetiva fiscalização do seu (in)adimplemento e sancionamento da sua violação.

PROENÇA/JOSÉ JOAQUIM FERNANDES OLIVEIRA MARTINS, *Igualdade de Género e Assédio na Magistratura Judicial Portuguesa (Uma primeira reflexão crítica)*, p. 38, https://julgar.pt/wp-content/uploads/2021/05/20210511-JULGAR-Igualdade-de-g%C3%A9nero-e-ass%C3%A9dio-Joaquim-Martins-Ana-Proen%C3%A7a.pdf).

26 Só outro exemplo, no Art. 4,º, n.º 2, dispõe-se que "*Os juízes dos Tribunais Judiciais não utilizam nenhuma informação confidencial a que tenham acesso por via das suas funções em benefício privado, próprio ou de terceiros", quando essa utilização pode até constituir um crime (rectius, até vários crimes) e é claramente um ilícito disciplinar (e não apenas uma questão de ética ou de falta dela), violando, antes, alguns dos deveres mais básicos e essenciais de qualquer pessoa que exerça funções num tribunal (e não apenas dos juízes)*".

27 Veja-se, a este respeito, o que escreve ANA CELESTE CATARRILHAS DA SILVA EVANS DE CARVALHO, *Um código de conduta para os juízes? Uma resposta (necessariamente) à luz da Constituição e da Lei*, pp. 36-37, https://julgar.pt/um-codigo-de-conduta-para-os-juizes-uma-resposta-necessariamente-a-luz-da-constituicao-e-da-lei/, "*Impondo a Constituição que as normas estatutárias, como as que respeitam aos deveres e obrigações do juiz e ao regime disciplinar, constem de um único estatuto e sejam aprovadas pela Assembleia da República, são concedidas aos Conselhos Superiores, enquanto órgãos de gestão e de disciplina dos juízes, as seguintes opções: (i) solicitar à Assembleia da República, diretamente ou através do Ministério da Justiça, a iniciativa legislativa de alteração do EMJ, de forma a alterar o regime dos deveres e incompatibilidades dos juízes e do seu respetivo regime disciplinar, procedendo ao seu alargamento, prevendo mais deveres e obrigações, ou (ii) adotar um Código de Conduta dos juízes, no exercício da sua competência administrativa, com a natureza de mera Recomendação, nele enunciando as máximas da ética profissional do juiz, tipificando as condutas desejavelmente a adotar e a não adotar, cujo desrespeito constitui uma mera falta ética ou deontológica, sem quaisquer repercussões disciplinares. Se assim não for, se nenhuma destas opções for prosseguida e a iniciativa dos Conselhos Superiores adotar forma ou conteúdo diferente, estaremos perante um regulamento administrativo não apenas ilegal, mas inconstitucional. Ilegal, com o duplo fundamento: (a.) de ser emanado por quem não dispõe da competência legal, incorrendo não apenas no vício de incompetência relativa, por incompetência do órgão, mas de incompetência absoluta, por falta de atribuições e, por isso, enfermado de nulidade, nos termos do disposto no artigo 161.º, n.ºs 1 e 2, a) e b) do CPA e (b.) por falta de forma, ao emanar um regulamento administrativo em matéria que tem de revestir a forma de lei. Inconstitucional, por violação do disposto no artigo 215.º, n.º 1 da CRP, ao prever num regulamento administrativo matéria de deveres e obrigações e seu regime disciplinar, que integra o regime estatutário dos juízes. Considerando a recente revisão do EMJ, ocorrida em 2019, que operou o alargamento dos deveres e obrigações do juízes, a melhor opção no atual contexto será a enunciada em (ii), de aprovar um Código de Conduta dos Juízes, sob a natureza jurídica de Recomendação, que sirva de padrão orientador da conduta desejavelmente a adotar pelos juízes, de cumprimento espontâneo e voluntário, decorrente da elevada ética profissional que preside aos titulares do poder judicial e cuja inobservância origina o cometimento de uma falta ética, do foro da deontologia profissional e não do regime do ilícito disciplinar*".

Efetivamente, esta aparente inação (temporal e de vinculatividade) do CSM nesta matéria contrasta com o empenho e até entusiasmo que vem sido demonstrado com o *Projeto Meenos*[28], destinado a efetuar "Recomendações linguísticas e de estilo"[29] (uma espécie, perdoe-se a expressão orwelliana, de novilíngua judicial) aos juízes (com, a meu ver e pelos motivos que poderei desenvolver um dia, duvidosa legalidade e até constitucionalidade[30]), pretendendo-se mobilizar a *longa manus* dos serviços de inspeção do CSM para a sua divulgação e imposição, por essa via, aos julgadores[31].

Deste modo, os riscos relativos ao uso da IA por juízes são bem evidentes e muito superiores aos que se pretendem acautelar com essas "recomendações linguísticas", dado que, como tenho sempre defendido, "*O recurso não autorizado e não transparente pelos julgadores, sem qualquer limitação ética, a sistemas de IA poderá minar o próprio sistema judicial e a sua legitimidade/legitimação, dado que a pergunta que se seguirá é saber se os 'intermediários' juízes não podem ser eliminados e substituídos totalmente por 'juízes-robôs' aparentemente mais eficazes e isentos*"[32].

Há, de facto, um claro risco[33] dos juízes (ou de uma parte relevante dos mesmos) assinarem um verdadeiro pacto faustiano (ou talvez "faustIAno"), "*vendendo a alma à IA*" e não respeitarem estas recomendações, limitando-se a descarregar em bruto os dados de um processo num sistema generalista de IA de acesso público e a pedir para lhes dar uma

28 https://csm.org.pt/tag/projeto-meenos/.

29 https://csm.org.pt/meenos-recomendacoes/.

30 E muito menos sentido num sistema judicial com defeitos e necessidade várias e que vão muito para lá da simples escrita e comunicação das decisões judiciais, num país em que as decisões são escritas, ao mesmo tempo e conforme o juiz, pré ou pós Acordo Ortográfico (ou, às vezes, até os dois) e em que a maioria dos arestos dos tribunais superiores são justamente o oposto do que aí é propugnado, não havendo também qualquer uniformização a nível gráfico e dos tipos de letras (*font*) a que se recorre.

31 MARGARIDA SOUSA, *Projeto Meenos – Para uma linguagem clara e eficaz na Justiça*, p. 25, https://csm.org.pt/wp-content/uploads/2025/04/2025_CSM_Boletim2.pdf, escreve que "*A deliberação do CSM conterá orientação no sentido da valorização da simplificação, clareza e compreensibilidade das decisões judiciais, por referência às recomendações e aos modelos acolhidos, nas futuras ações inspetivas. A divulgação dos aludidos textos de apoio e modelos será efetuada pelo CSM e os inspetores farão ações de sensibilização para a sua utilização.* Dois anos após a divulgação dos modelos acolhidos, o CSM procederá à avaliação dos resultados alcançados" (itálico acrescentado).

32 JOSÉ JOAQUIM FERNANDES OLIVEIRA MARTINS, «Legitimidade, legitimação e utilização de ferramentas informáticas (em particular de sistemas de Inteligência Artificial – IA) pelos juízes» in *Inteligencia Artificial en la Justicia y en el Arbitraje: discusiones y perspectivas*, p. 51, https://www.tprmercosur.org/es/docum/cmped/libro2026_ia.pdf.

33 Sendo que já existem, há algum tempo, rumores que tal já terá sucedido (e ainda poderá ocorrer entre nós) e que há até quem se gabe publicamente desse tipo de atuações perante os mandatários das partes.

sentença, usando, sem qualquer revisão e controlo, o resultado obtido, por muito mau que seja (e por mais injusta que seja a respetiva decisão final).

Se o uso intensivo da IA pode conduzir à perda das competências mais básicas dos juízes (o já referido *deskilling*) ou ao aceitar e reproduzir acriticamente e sem qualquer ponderação e reflexão todo e qualquer resultado dos sistemas de IA (efeito de ancoragem, um bem conhecido viés cognitivo), limitando-se o julgador a ser um simples, surdo e mudo, núncio desse sistema, afigura-se que será importante que estas recomendações sejam mais do que isso, sejam antes regras vinculativas e correspondam a deveres efetivos dos juízes, objeto de fiscalização e passíveis de sancionamento[34].

Aliás, mais do que isso, o ideal seria que todo o sistema informático passasse, como já se disse, para o controlo dos Conselhos Superiores e que as aplicações informáticas utilizadas não possibilitassem a exportação dos dados dos processos judiciais para utilizações indevidas, nomeadamente em sistemas de IA não autorizados, e só permitissem que os julgadores recorressem exclusivamente aos sistemas de IA auxiliares certificados e admitidos para o efeito (com formação dos julgadores para a forma como devem - ou não - utilizar os mesmos), evitando uma possível e muito prejudicial "corrida à IA" só para reduzir o trabalho judicial e premiada com uma melhor classificação dos julgadores que não cumpram estas recomendações.

Voltando à música e em jeito de uma *coda* final, cumpre reiterar que só com uma intervenção atempada, assertiva e adequada dos Conselhos Superiores, passando estas recomendações a verdadeiras e efetivas regras jurídicas, vinculativas e passíveis de fiscalização e sancionamento, se poderá concluir que "*será lícita a utilização de sistemas de IA de apoio à decisão judicial [judiciária], após certificação e quando não se trate de uma "prática de IA proibida", sempre e quando sejam sistemas lógico-simbólicos estruturados exclusiva e robustamente por programadores humanos, salvo se com uma relevância secundária, suscetível de afastar a aplicabilidade dos regime relativo aos "sistemas de inteligência artificial de risco elevado" (Art.º 6.º n.º 3), ou - se tais sistemas vierem a cumprir os*

34 E que se reflitam, aliás, na própria classificação dos juízes, para o que será necessário que o seu (in)cumprimento possa ser verificado e apreciado pelos inspetores judiciais (que deverão, claro, ter formação para o efeito), dado que tal repercussão a nível da notação será sempre bem mais importante do que encarregar os inspetores de divulgar e aferir do cumprimento de regras estilísticas e até 'coagir', pela sua repercussão a nível classificatório, à submissão a essas normas linguísticas, numa espécie de 'jogos florais' destinados a premiar os juízes mais permeáveis e obedientes aos ditames linguísticos do CSM e dos seus consultores.

requisitos exigidos para os "modelos de IA de finalidade geral com risco sistémico", atendendo aos "efeitos negativos reais ou razoavelmente previsíveis [dos "modelos de IA de finalidade geral"] nos direitos fundamentais" (Art.ºs 3.º 65) e 51.º a 56.º, esclarecidos nos Considerandos (63) a (66), (97) e (110) a (120))"[35].

35 MANUEL DAVID MASSENO, *Da Justiça no Regulamento da IA da União Europeia*, p. 7, https://www.academia.edu/130350079/DA_Justi%C3%A7a_no_Regulamento_da_IA_da_Uni%C3%A3o_Europeia.



CYBERLAW

BY CIJIC

IUS EX MACHINA? ENTRE UMA ADMINISTRAÇÃO PÚBLICA "*ARTIFICIALMENTE INTELIGENTE*" E A TUTELA DOS DIREITOS FUNDAMENTAIS DOS ADMINISTRADOS

LUÍS MANUEL PICA*

* Doutor em Ciências Jurídicas Públicas pela Escola de Direito da Universidade do Minho; Professor Adjunto do Instituto Politécnico de Beja; Professor Auxiliar Convidado do Instituto Superior Manuel Teixeira Gomes; Investigador Doutorado-Integrado do CEAD - Francisco Suárez da ULHT e Investigador do JusGov - Centro de Investigação em Governação e Justiça da EDUM; https://orcid.org/0000-0002-0470-5732; luis.pica@ipbeja.pt

RESUMO:

A expansão da Inteligência Artificial na Administração Pública desafia a dogmática clássica do Direito Administrativo, exigindo a conciliação da eficiência tecnológica com as garantias do Estado de Direito. A procedimentalização algorítmica transforma a tramitação procedimental num ecossistema híbrido, diluindo a vontade humana e fragmentando a autoria do ato. Perante este paradigma, torna-se necessário revisitar os institutos clássicos do Direito Administrativos, os quais se consideravam plenamente sedimentados, a fim de dar resposta aos novos desafios trazidos pela implementação da inteligência artificial na atividade administrativa.

Palavras-chave: Administração Pública; Inteligência Artificial; Procedimento Administrativo Algorítmico; Decisão Automatizada; Garantias Procedimentais.

ABSTRACT:

The expansion of Artificial Intelligence within Public Administration fundamentally challenges the classical dogmatics of Administrative Law, demanding a harmonious reconciliation between technological efficiency and the core guarantees of the Rule of Law. Algorithmic proceduralizing transforms standard administrative proceedings into a hybrid ecosystem, dissolving human volitional intent and fragmenting the attribution of authorship. Faced with this paradigm, it becomes necessary to revisit classical administrative law institutes - previously considered fully settled - to address the new challenges brought by the implementation of artificial intelligence in administrative activity.

Keywords: Public Administration; Artificial Intelligence; Algorithmic Proceduralizing; Human-in-the-loop; Enhanced Justification; Institutional Authorship.

I. INTRODUÇÃO

A crescente integração da Inteligência Artificial na atividade das Administrações Públicas está a operar uma transformação ontológica no procedimento administrativo, transfigurando o modelo tradicional de decisão humana num ecossistema tecnologicamente mediado.

Sob a égide da eficiência computacional, assiste-se à emergência de um procedimento administrativo cada vez mais dependente de sistemas de decisão artificialmente inteligente, fenómeno em que os algoritmos e o código-fonte assumem uma função materialmente normativa, redefinindo as dinâmicas de poder e as garantias dos cidadãos.

Esta mudança paradigmática coloca severos desafios à dogmática clássica do Direito Administrativo, concebida para um modelo decisório fundado no discernimento humano, na razão discursiva e no diálogo procedimental.

Sob esta problemática, o presente artigo procura analisar e dar um contributo (pequeno) sobre o impacto que estes sistemas preditivos e de decisão tem no procedimento administrativos, e consequentemente sobre as estruturas constitucionais do Estado de Direito, com particular incidência na erosão do dever de fundamentação diante da opacidade algorítmica (*black box*) e no risco de supressão fática da audiência prévia. Defende-se a urgente reconfiguração das categorias dogmáticas tradicionais - da discricionariedade aos vícios de validade do ato -, propondo-se uma reserva de decisão humana para atos não vinculados e a consagração de um dever de fundamentação algorítmica reforçada. Em última análise, pretende-se demonstrar que a eficiência tecnológica só encontra legitimidade democrática se for estritamente subordinada aos princípios constitucionais da legalidade, igualdade, imparcialidade e transparência.

II. O CONCEITO JURÍDICO DE INTELIGÊNCIA ARTIFICIAL: DA AUTOMATIZAÇÃO À COGNIÇÃO COMPUTACIONAL

Enquadramento

Falar em digitalização e sistemas de inteligência artificial parece, numa primeira aproximação, abordar as realidades perante um literalismo homogéneo e tendencialmente uniforme, pois parecem realidades análogos, sendo, por vezes, utilizados indistintamente[1].

Por exemplo, se tomarmos como exemplo o disposto no n.º1 do artigo 14.º do Código do Procedimento Administrativo, o qual dispõe que, "*os órgãos e serviços da Administração Pública devem utilizar meios eletrónicos no desempenho da sua atividade, de modo a promover a eficiência e a transparência administrativas e a proximidade com os interessados",* parece que o conceito de "meios eletrónicos" abrange qualquer realidade, sendo utilizado como um termo com enorme amplitude semântica. Ou seja, do mencionado normativo não se vislumbra uma cisão entre os mencionados conceitos de digitalização e automatização. A lei refere, apenas, a utilização de meios eletrónicos, o qual parece apontar para qualquer uma das aludidas realidades[2].

Porém, tal ideia não poderá deixar de revestir um certo grau de confusão e, até, tendencial engano, já que os conceitos de digitalização e automatização são realidades distintas, com implicações diferentes e repercussões que são profundamente heterogéneas:

a) De forma mais precisa, podemos entender a *digitalização* como a transposição do procedimento administrativo e tributário tradicional para um suporte digital, sem alterar a essência do mesmo. Trata-se de um modelo que utiliza plataformas eletrónicas em vez de documentos físicos, mantendo inalteradas as fases, os atores e os atos envolvidos no procedimento. A única diferença reside na substituição do suporte utilizado por um formato digital.

b) Por outro lado, a *automatização* traz um *plus,* pois existe uma transformação mais profunda, introduzindo sistemas de apoio à decisão, como a inteligência

1 LUÍS MANUEL PICA, "*Digitalização, Inteligência Artificial e controlo jurisdicional da atividade administrativa tributária*", *in Comentários à Legislação Processual Tributária.* Lisboa, AAFDL Editora, 2025.
2 Entre outros, veja-se o nosso, LUÍS MANUEL PICA, *Inteligência Artificial no Direito Administrativo - Tensões, Problemáticas e Soluções dos Institutos Clássicos*, Coimbra, Almedina, 2026, pp.13 e ss.

artificial, que podem substituir os agentes humanos na tomada de decisões e na execução dos atos procedimentais. Aqui, vai-se além da simples digitalização, utilizando critérios pré-estabelecidos para interpretar e aplicar as normas jurídicas, formulando projetos de decisões, ideias ou conclusões que podem ser adotadas no procedimento[3]. A inovação reside não na decisão em si, mas no processo que a antecede: a decisão inteligente é o resultado da análise e processamento de dados pela máquina, enquanto a decisão digital se limita à informatização do procedimento já existente.

Na situação descrita (da *automatização*), colocam-se, contudo, três realidades distintas, mas necessárias entre si:

a) a realidade em que os dados - no nosso caso, os dados públicos - implicam o elemento estruturante de um novo "modelo" de Administração - uma "*Administração Datificada*" - em que os programas farão o que os burocratas faziam, funcionando como um sistema puramente reativo de "entrada-decisão"[4];

b) a realidade na qual a modernização, a automatização e a integração dos sistemas de inteligência artificial redefinem a marca do Direito Administrativo como direito positivo, porque esses processos exigem a tradução da linguagem natural "típica" da norma jurídica para uma linguagem "informatizada";

c) a tomada de decisões é conseguida com suporte a um elemento externo que forma um elemento persuasivo na decisão final, podendo, inclusive, colocar-se em crise a vontade da administração.

3 Neste sentido, veja-se a definição dada pelo 3.º, n.º1, do Regulamento de Inteligência Artificial, aprovado pelo Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, o qual estabelece que um Sistema de Inteligência Artificial será "*um sistema baseado em máquinas concebido para funcionar com níveis de autonomia variáveis, e que pode apresentar capacidade de adaptação após a implantação e que, para objetivos explícitos ou implícitos, e com base nos dados de entrada que recebe, infere a forma de gerar resultados, tais como previsões, conteúdos, recomendações ou decisões que podem influenciar ambientes físicos ou virtuais"*.

4 LUÍS MANUEL PICA, *A inteligência artificial e o Direito Tributário: Fundamentos e Limites Constitucionais,* Coimbra, Almedina, 2023, p.119.

Por isso, não podemos deixar de concluir pela necessidade de existir uma cisão dos conceitos. A adoção de sistemas inteligentes carece de uma norma legitimadora que permita à administração a utilização destes sistemas de novas tecnologias e de meios de automatização e de apoio à decisão.

Breve contributo para uma noção (aproximada) de Inteligência Artificial

A definição de inteligência artificial constitui uma das questões mais controversas na interface entre direito e tecnologia. A doutrina tem procurado distinguir entre a mera automatização - que implica execução mecânica de instruções humanas - e a inteligência artificial em sentido forte, caracterizada por autonomia cognitiva, capacidade de aprendizagem e interação com outros sistemas.

Em termos técnicos e jurídicos, pode afirmar-se que a inteligência artificial é um sistema computacional dotado de capacidade de aprender a partir de dados, interagir com o meio digital e cooperar com outros agentes, humanos ou artificiais, reformulando autonomamente os próprios modelos de decisão. Esta conceção, que ultrapassa a simples automação, insere-se naquilo que a doutrina italiana designa de "*cognizione computazionale*", ou seja, *um processo de aquisição e desenvolvimento de conhecimento que se aproxima das formas humanas de raciocínio, embora sem consciência ou intencionalidade*[5].

Enquanto a automação repete regras predefinidas, a inteligência artificial modifica-as, gera-as e aplica-as em novos contextos revelando uma dimensão criativa e adaptativa. A diferença é, portanto, qualitativa, já que *o automatismo é reativo,* enquanto a inteligência artificial *é adaptativa.* A inteligência artificial não é, aqui, mais um mero instrumento técnico, mas um sujeito funcional da ação administrativa, dotado de autonomia operacional dentro de limites juridicamente determinados. Poderá dizer-se que se enquadra enquanto "sujeito funcional" (*soggetto funzionale*), que não é pessoa jurídica, mas atua como extensão cognitiva da Administração, desempenhando tarefas de análise, previsão e decisão dentro de parâmetros programados e supervisionados. Ou seja, a inteligência artificial funciona como infraestrutura de raciocínio coletivo, que apoia ou substitui parcialmente a racionalidade humana na tomada de decisão pública.

5 LUÍS MANUEL PICA, *Inteligência Artificial no Direito Administrativo - Tensões, Problemáticas e Soluções dos Institutos Clássicos*, Coimbra, Almedina, 2026, pp.13 e ss.

A inteligência artificial constitui um novo nível de racionalidade delegada, no qual o decisor não abdica da sua função, mas a amplia por meio de ferramentas capazes de aprender e dialogar com outros sistemas. Assim, a inteligência artificial é compreendida como um sistema interativo, capaz de integrar dados de múltiplas bases e de cooperar com outras inteligências artificiais em ecossistemas administrativos complexos.

Porém, a inteligência artificial não é um autómato mecânico, mas uma entidade dinâmica que aprende, evolui e co-constrói decisões com o homem. A inteligência artificial representa um novo paradigma de co-decisão homem-máquina, onde a racionalidade administrativa se torna híbrida, exigindo novos instrumentos de imputação e responsabilidade, pois *a máquina é capaz de interagir com outras máquinas, com o ambiente e com a sociedade, gerando comportamentos não inteiramente previsíveis pelo programador*[6].

Esta imprevisibilidade não é defeito, mas elemento essencial do seu potencial de aprendizagem. O Direito, todavia, não pode ignorar que essa autonomia técnica exige mecanismos de contenção jurídica, de modo a evitar que a capacidade de autoaperfeiçoamento comprometa o princípio da legalidade.

Deste modo, dir-se-á que a inteligência artificial tende, portanto, a conceber-se como entidade cognitiva relacional, dotada de capacidade para:

1. Interagir com dados, utilizadores e outras inteligências artificiais;
2. Aprender autonomamente com base em experiência e retroalimentação (*feedback loops*);
3. Adaptar os próprios modelos de decisão a novos contextos;
4. Evoluir intelectualmente, mediante reorganização interna das suas regras heurísticas;
5. Contribuir funcionalmente para a formação da vontade administrativa, sem substituí-la.

Esta caracterização é consistente com a definição adotada no *AI Act*, aprovado pelo Regulamento (UE) 2024/1689, onde se lê no seu artigo 3.º, n.º 1, que os sistemas de inteligência artificial são "*sistemas baseados em algoritmos que, para um determinado conjunto de objetivos definidos pelo ser humano, geram saídas como previsões,*

6 LUÍS MANUEL PICA, *Inteligência Artificial no Direito Administrativo - Tensões, Problemáticas e Soluções dos Institutos Clássicos*, Coimbra, Almedina, 2026, pp.13 e ss.

recomendações ou decisões que influenciam ambientes físicos ou virtuais". O legislador europeu reconhece, portanto, o carácter adaptativo e autónomo desses sistemas, embora sublinhe a necessidade de controlo humano significativo.

A definição normativa é um ponto de partida, mas deve-se partir mais além do texto normativo. Isto porque, a inteligência artificial interpreta não como simples tecnologia, mas como nova forma de racionalidade jurídica operacional e como elemento da ação administrativa com dignidade epistemológica própria. Por isso, o desafio é construir um Direito capaz de compreender as máquinas que aprendem, isto é, capaz de integrar a lógica evolutiva da inteligência artificial nos quadros da legalidade e da responsabilidade.

III. ADMINISTRAÇÃO ELETRÓNICA *VERSUS* ADMINISTRAÇÃO ALGORÍTMICA

Enquadramento

A passagem da *administração eletrónica* (caracterizada pela desmaterialização documental e pela automação de tarefas rotineiras) para a *administração algorítmica* (na qual as decisões substantivas são delegadas em sistemas inteligentes) implica uma mudança qualitativa. Estamos perante uma "autonomia funcional da tecnologia" que desafia os conceitos tradicionais de autoria e de responsabilidade administrativa. Enquanto a administração eletrónica mantém o núcleo da decisão humana, a administração algorítmica tende a transferir a ponderação e a decisão para sistemas de inteligência artificial, os quais podem sofrer de opacidade e de difícil auditabilidade, com influência de fatores externos e de outros sistemas de inteligência artificial que possam intervir com a utilizada pela administração. Esta transformação torna difusa a fronteira entre o procedimento administrativo e o processo de cálculo automatizado, o que gera riscos de vícios de forma, de fundamentação e de imparcialidade, como se explorará nos pontos seguintes.

A natureza da administração eletrónica

Durante algumas décadas, acreditou-se que a "*globalização*" era um fator decisivo no desenvolvimento das relações sociais, económicas, políticas e jurídicas[7].

A ideia difundida era de que a cooperação entre pessoas, entidades e pelos próprios Estados, juntamente com a abertura dos mercados e das relações, estimulava a competitividade em diversas esferas do quotidiano, gerando uma necessidade constante de evolução para acompanhar a competitividade e as demandas evolutivas. E, embora essa visão tenha fundo de verdade, constata-se que a globalização é um fenómeno que gera competitividade e combate o isolamento, sugerindo que o ser humano é impulsionado a se desenvolver para não ficar afastado em relação aos outros.

Mas, para além disto, é essencial reconhecer que as crises sociais são um fator de impulso, designadamente gerando uma necessidade de mudança. Em tempos de crise, a complexidade surge como um obstáculo nas relações sociais, económicas e, também, jurídicas, afetando profundamente o Direito, assim como outras áreas da ordem social. E, neste sentido, o Direito Administrativo, naturalmente, não é exceção. Exemplos disso incluem as crises económicas do final da primeira década do século XXI, a pandemia de COVID-19, as guerras que se iniciaram em 2022 na Ucrânia e mais recentemente no Oriente Médio, e as crises humanas, materiais e económicas resultantes dessas situações.

Todavia, essas implicações não precisam ser inteiramente negativas, pois as crises ocorridas, apesar da catástrofe global que representaram, aceleraram exponencialmente os processos de transformação necessários e evidentes, apresentando uma grande oportunidade para o aprimoramento da Administração Pública. Por essa razão, discute-se, no âmbito da atividade administrativa e da organização da Administração Pública, a ideia de uma *administração eletrónica*[8]. É que, o advento das novas tecnologias deu origem a uma revolução tecnológica, com inegáveis consequências no domínio económico e social que está a afetar a administração e, com ela, a própria forma do Direito Administrativo. Não falamos de um novo ramo do Direito, ou de uma nova forma de organização administrativa ou

7 LUÍS MANUEL PICA, *Inteligência Artificial no Direito Administrativo - Tensões, Problemáticas e Soluções dos Institutos Clássicos*, Coimbra, Almedina, 2026, pp.13 e ss.

8 MIGUEL PRATA ROQUE, "*O Nascimento da Administração Eletrónica num Espaço Transnacional (Breves notas a propósito do projeto de revisão do Código do Procedimento Administrativo)*", *in e-Pública* - Vol. I Nº. 1, janeiro 2014, pp.307-321

procedimental. O que define uma *administração eletrónica* é a utilização das tecnologias de informação e comunicação nas administrações públicas, combinada com mudanças organizacionais e novas competências, a fim de melhorar os serviços públicos e os processos democráticos e de reforçar o apoio às políticas públicas. Trata-se de uma realidade que não é despicienda no ordenamento jurídico português, pois veja-se que o artigo 60.º-A, n.º1, da Lei Geral Tributária e o artigo 14.º do Código do Procedimento Administrativo já preveem a utilização das novas tecnologias no procedimento tributário[9].

Com efeito, a Administração Pública eletrónica não é mais do que a combinação que os sistemas de tecnologia de informação e comunicação com a administração, procurando atingir o objetivo de melhorar a satisfação do interesse público. Não introduz um modelo de administração que rompa radicalmente com o anterior. A administração continua a ser uma administração, uma organização pública privilegiada cujo objetivo é a satisfação do interesse geral. O que muda são as suas formas: tanto a organização administrativa como o exercício da função administrativa são marcados pela utilização das novas tecnologias, com tudo o que isso significa[10]. Do mesmo modo, o procedimento administrativo eletrónico não é diferente do procedimento administrativo comum, mas sim o resultado da aplicação das TIC a este último, com as novas possibilidades de melhor cumprimento do princípio da eficiência que daí advêm[11].

A natureza do ato administrativo emanado pelo sistema de Inteligência Artificial

Conforme se explicou, a digitalização e automação do procedimento administrativo são uma possibilidade que altera arquétipo clássico da atividade procedimental, levando-nos a questionar sobre a natureza do(s) ato(s) administrativo(s) praticado(s). Isto porque, a incorporação de meios eletrónicos - seja por via da mera digitalização do procedimento tributário ou pela implementação de sistemas de inteligência artificial - suscita uma das questões mais acutilantes no "Direito Administrativo Digital"[12], nomeadamente saber se o

9 LUÍS MANUEL PICA, "*Digitalização, Inteligência Artificial e controlo jurisdicional da atividade administrativa tributária*", *in Comentários à Legislação Processual Tributária.* Lisboa, AAFDL Editora, 2025.
10 SANTI ROMANO, *Principii di Diritto Amministrativo Italiano*, Milão, CEDAM, 1937, p.77.
11 MARCELO REBELO DE SOUSA, ANDRÉ SALGADO DE MATOS, *Direito Administrativo Geral - Tomo I - Direito Administrativo Geral. Introdução e Princípios Fundamentais*, 3ª Edição, Dom Quixote, 2008, pp.38 e seguintes.
12 JEAN-BERNARD AUBY, "Administrative Law Facing Digital Challenges", *in European Review of Digital Administration & Law - Erdal 2020,* Volume 1, Issue 1-2, June-December, pp.7-15, disponível em *https://www.erdalreview.eu/free-download/97888255389602.pdf* [consultado em 13.07.2026].

ato administrativo eletrónico é efetivamente um ato tributário, na conceção dogmática e jurídica estabelecida em torno do Direito Administrativo clássico.

Como a Administração é uma organização voltada para a realização do interesse público, torna-se essencial definir a natureza dos atos que ela pratica com o suporte de meios eletrónicos. Estabelece-se que, a questão em torno do ato tributário mostra-se pertinente, pois não é possível abordar seriamente a temática subjacente ao controlo jurisdicional dos atos, enquanto forma de fiscalização da atividade levada a cabo pela administração, sem antes fazermos uma breve alusão ao ato administrativo digital, ou, numa amplitude maior, ao ato administrativo baseado num procedimento totalmente automatizado[13].

Do ponto de vista concetual, partimos da noção normativa dada ao ato administrativo, considerando o disposto no artigo 148.º do Código do Procedimento Administrativo. Consagra o citado normativo legal que, "*consideram-se atos administrativos as decisões que, no exercício de poderes jurídico-administrativos, visem produzir efeitos jurídicos externos numa situação individual e concreta"* (sublinhado nosso)[14]. Em matéria tributária, a noção adotada não poderá ser distinta, apesar da legislação em matéria tributária não consagrar qualquer definição para efeitos legais, pelo que, não havendo qualquer uma consagração normativa, caberá à doutrina tal incumbência. Para MÁRIO ESTEVES DE OLIVEIRA, PEDRO COSTA GONÇALVES e J. PACHECO DE AMORIM[15], o ato administrativo assume as seguintes caraterísticas:

> *"a) Consistir numa decisão, enquanto estatuição ou prescrição, voluntária;*
>
> *b) Essa decisão ser proferida por órgãos ou agentes da administração no exercício de poderes e deveres de autoridade administrativa;*
>
> *c) Essa decisão ter por base normas de direito público, isto é, normas de competência que regulam situações e relações jurídicas que pelo seu sujeito e conteúdo são insuscetíveis de se constituir entre simples particulares;*

13 Neste sentido, refere DIOGO FREITAS DO AMARAL que, "originariamente, o conceito de acto administrativo foi construído com base em considerações de natureza jurisdicional: ele delimitava certos comportamentos da Administração em função da fiscalização da actividade administrativa pelos tribunais". DIOGO FREITAS DO AMARAL *Curso de Direito Administrativo - Volume II,* 2ª Edição, Coimbra, Almedina, 2011, p.231.

14 Sobre o conceito de ato administrativo, VASCO PEREIRA DA SILVA, *Em busca do Ato Administrativo Perdido,* Coimbra, Almedina, 2016; DIOGO FREITAS DO AMARAL *Curso de Direito Administrativo - Volume II,* 2ª Edição, Coimbra, Almedina, 2011, pp.236 e seguintes.

15 MÁRIO ESTEVES DE OLIVEIRA, PEDRO COSTA GONÇALVES e J. PACHECO DE AMORIM, *Código do Procedimento Administrativo Comentado*, 2ª Edição, Coimbra, Almedina, 2003, p.550.

d) Produzir efeitos externos, na medida em que se produzem na esfera jurídica de terceiros que com o autor do ato estão, pretendam ou possam estar em relação jurídico-administrativa; e

e) Incidir numa situação individual e concreta".

No contexto da integração de sistemas de inteligência artificial no procedimento administrativo, um ato administrativo praticado integralmente por meios eletrónicos constitui uma declaração de vontade emitida pelo órgão competente com autoridade normativa para tal. O que muda, é o seu suporte físico. Um ato administrativo eletrónico não deixa de ser um ato volitivo da Administração destinado, por exemplo, à liquidação de um tributo, uma declaração de julgamento (*v.g.* apreciação da legalidade de um ato tributário em sede de reclamação graciosa), ou uma declaração de conhecimento (*v.g.* ato de reconhecimento de um benefício fiscal). Por isso, o ato administrativo emanado por meios eletrónicos assume o mesmo conteúdo que um ato precedido de um procedimento administrativo clássico emanado através de suportes físicos.

Da mesma forma, o facto de a decisão ter sido emanada por uma máquina de forma automatizada não implica que a ação lhe seja imputada, pois o utilizador deste sistema de inteligência artificial é sobre quem vai recair a autoria do ato praticado, pelo que a autoria recai sobre o órgão administrativo que exerce o poder e ao qual é confiada a competência normativa para a prática daquele ato[16]. Veja-se que a imputação de um ato não é mais do que um ato fictício, pelo que, a autoria de uma ação realizada por uma pessoa - ou, neste caso, por uma aplicação informática - integrada num órgão de uma administração pública é atribuída a esta última porque é imposta por lei. O elo entre a execução material da ação e o órgão administrativo com competência legal para a prática do ato é o dirigente desse mesmo órgão administrativo a quem foi atribuída a competência exercida. E, atualmente, esse titular é sempre uma pessoa.

Para além disso, devem ter cumprido dois requisitos: *a nomeação (livre ou através de um processo seletivo) e a tomada de posse.* A máquina –ferramenta informática - não é mais do que um meio instrumental, como tantos outros, ao serviço do chefe do órgão. A diferença,

16 MARCELO REBELO DE SOUSA, ANDRÉ SALGADO DE MATOS, *Direito Administrativo Geral - Tomo III - Actividade administrativa*, 2ª Edição, Dom Quixote, 2009.

neste caso, reside no facto de o meio material não se limitar a ajudar o dirigente do órgão a tomar uma decisão no exercício das suas funções, mas proceder diretamente à tomada de decisão, sem qualquer intervenção do dirigente do órgão nesse caso concreto. A ligação entre a decisão baseada em sistemas inteligentes e o titular do órgão ocorrerá através da assinatura eletrónica do chefe do órgão com competência para tal[17]. E, será por via desta assinatura que será imputado o ato ao órgão administrativo que representa.

Em síntese conclusiva, não podemos deixar de assumir que os atos praticados pela administração pública, mesmo que sustentados em sistemas de inteligência artificial, assumem a natureza de verdadeiros atos administrativos, pois são praticados por órgãos com competência normativa para a sua prática, incluídos num procedimento tributário normativamente aprovado pelo órgão legiferante competente, unilateral e produtor de efeitos jurídicos na esfera jurídica de um contribuinte. O facto de o ato ser praticado por sistemas de inteligência artificial não lhe retira estas caraterísticas, pois, por um lado, a supervisão humana é necessária para a validação das decisões tomadas, sendo esta supervisão obrigatoriamente feita por pessoas singulares que integram o órgão administrativo, devendo ser estes o verdadeiro elo entre o sistema de inteligência artificial e o órgão administrativo com competência para a prática do ato.

IV. OS RISCOS DA ADMINISTRAÇÃO PÚBLICA "*ARTIFICIALMENTE INTELIGENTE*" NA DOGMÁTICA CLÁSSICA DOS VÍCIOS PROCEDIMENTAIS ADMINISTRATIVOS

Ponto de partida

O procedimento administrativo, tradicionalmente compreendido como uma sequência ordenada e encadeada de atos preparatórios e complementares destinados à formação e exteriorização da vontade da Administração Pública, desempenha uma dupla função no Estado de Direito democrático: por um lado, visa assegurar a conformidade da atuação administrativa com a legalidade objetiva; por outro, tutelar as posições jurídicas subjetivas dos administrados. A dogmática clássica erigiu as patologias procedimentais - com especial

17 Sobre isto, Cfr. o nosso "*Digitalização, Inteligência Artificial e controlo jurisdicional da atividade administrativa tributária*", *in Comentários à Legislação Processual Tributária.* Lisboa, AAFDL Editora, 2025, pp.206-207.

destaque para a preterição de formalidades essenciais, a violação das regras de competência e a preterição da audiência dos interessados - em vícios geradores de invalidade jurídica, precisamente porque tais omissões comprometem a racionalidade discursiva, a transparência e a legitimidade democrática da decisão final.

Todavia, o surgimento de meios digitais e a incorporação de sistemas de Inteligência Artificial no seio da atividade administrativa provocam uma mutação ontológica na própria estrutura do procedimento.

À medida que a Administração substitui tarefas procedimentais por operações automatizadas - como o cruzamento massivo de bases de dados, a análise preditiva de risco fiscal ou o deferimento automático de prestações sociais -, opera-se o fenómeno da *"procedimentalização algorítmica".*

Esta transformação gera severas fraturas na dogmática clássica do Direito Administrativo. Num *plano ontológico***,** verifica-se a diluição das fronteiras entre a intenção humana voluntária e a execução mecânico-estatística. O conceito clássico de ato administrativo como "*manifestação de vontade de um órgão administrativo*" é colocado em crise quando o conteúdo declaratório resulta, direta ou predominantemente, de uma inferência computacional; num *plano epistémico,* o raciocínio jurídico-discursivo - assente na interpretação hermenêutica da norma e na valoração dos factos - é progressivamente substituído por um raciocínio estatístico e correlacional. O algoritmo não compreende nem interpreta os factos no sentido hermenêutico tradicional, limitando-se a processar padrões de dados e a calcular probabilidades; por último, num *plano axiológico***,** o eixo de legitimação do procedimento desloca-se da razão pública e da participação democrática para os valores da eficiência técnica, da celeridade procedimental e do rendimento computacional.

Tal mutação conduz inevitavelmente à *fragmentação da autoria procedimental* e à desmaterialização da vontade administrativa. Torna-se difuso determinar em que momento, e por quem, foi efetivamente formada a vontade decisória, gerando uma opacidade estrutural que desafia o paradigma do Código do Procedimento Administrativo.

Vejamos, assim, os fundamentos dogmáticos que devem ser acautelados perante os riscos da integração da inteligência artificial no seio da Administração Pública.

O exercício do poder interpretativo perante a elasticidade semântica da lei

A adoção de conceitos jurídicos indeterminados (ou tendencialmente abertos) e a concessão de margem de livre apreciação à Administração Pública constituem mecanismos indispensáveis de adaptação da lei à complexidade da vida social e económica. Termos como "interesse público", "manifesta urgência", "injustiça grave" ou "relevância social" não configuram meras cláusulas estilo, pois encerram uma elasticidade semântica que exige dos órgãos administrativos uma apreciação valorativa, contextualizada e casuística.

A elasticidade semântica não equivale a um poder arbitrário ou à ausência de direito. Ou seja, a elasticidade representa um espaço de valoração juridicamente delimitado, balizado de forma estrita pelos princípios constitucionais da legalidade, da proporcionalidade, da igualdade, da imparcialidade, da boa-fé e da boa-administração (cfr. artigos 266.º, n.º 1 e 2 da Constituição da República Portuguesa e artigo 3.º, n.º 1 do Código do Procedimento Administrativo).

Coloca-se, pois, a questão sobre se poderá um sistema de inteligência artificial exercer, autonomamente, valorações interpretativas ou preencher conceitos tendencialmente abertos que exijam certa ponderação de interesses?

A resposta dogmática deve ser *perentoriamente negativa*, pois o raciocínio algorítmico é de matriz determinística, lógica ou estocástica, isto é, baseia-se em regras pré-programadas ou em otimizações estatísticas extraídas de dados ou acontecimentos do passado. O algoritmo é incapaz de realizar valorações ético-jurídicas no presente ou de valorar a equidade de uma situação atípica que fuja aos padrões estatísticos do seu modelo de treino. Como referencial paradigmático de regulação, aponta-se, por exemplo, a solução proposta e inspirada no § 35a da Lei de Procedimento Administrativo Alemã *(Verwaltungsverfahrensgesetz - VwVfG)*[18]. O referido preceito estabelece que um ato administrativo apenas pode ser emitido de forma totalmente automatizada se concorrerem três requisitos cumulativos:

1. Existência de uma *habilitação legal expressa* que autorize a automação daquele procedimento específico;

18 Proposta já apresentada por nós em: LUÍS MANUEL PICA, *Inteligência Artificial no Direito Administrativo - Tensões, Problemáticas e Soluções dos Institutos Clássicos*, Coimbra, Almedina, 2026, pp.69 e ss.

2. Tratar-se de uma atuação estritamente vinculada, na qual *não exista margem de discricionariedade* nem margem de livre apreciação por parte da autoridade pública;
3. Cumprimento de *requisitos técnicos de auditabilidade e monitorização humana*.

Transpondo esta construção para o ordenamento jurídico português, conclui-se que a decisão administrativa totalmente automatizada é manifestamente incompatível com o Estado de Direito material sempre que a norma aplicável exigir a ponderação casuística de interesses ou o preenchimento de conceitos indeterminados. Nestas hipóteses, impõe-se uma rigorosa *reserva de decisão humana*, sob pena de usurpação do poder valorativo que a lei confiou única e exclusivamente à autoridade pública dotada de discernimento e de legitimidade democrática.

O dever constitucional de fundamentação e o desafio do Machine Learning

O dever de fundamentação dos atos administrativos que afetem direitos ou interesses legalmente protegidos consagra-se como um pilar de ordem pública constitucional (artigo 268.º, n.º 3 da Constituição da República Portuguesa), traspondo uma preocupação outrora vivida de opacidade administrativa e decisória no âmbito da atividade procedimental (artigo 152.º do Código do Procedimento Administrativo). A fundamentação cumpre uma tríplice função:

1) Garantir a autodisciplina e a ponderação racional da Administração no momento da decisão;
2) Permitir ao administrado o conhecimento do percurso cognoscitivo e valorativo (*iter cognitivo*) para que possa optar conscientemente entre a aceitação do ato ou a sua impugnação contenciosa;
3) Viabilizar o controlo jurisdicional pelos órgãos jurisdicionais competentes[19].

19 No Ac. do STA de 19/2/2003 é referido que "*a fundamentação dos atos administrativos visa, por um lado, dar a conhecer aos seus destinatários o iter cognoscitivo e valorativo seguido pela Administração, de molde a permitir-lhes uma opção consciente entre a aceitação do ato e a sua impugnação contenciosa, e, por outro, que a Administração, ao ter de dizer a forma com agiu, porque decidiu desse modo e não de outro, tenha de ponderar aceitavelmente a sua decisão*". Ac. do STA de 19/2/2003, proc. n.º 762/02, relator Conselheiro António Madureira, disponível em http://www.dgsi.pt (consultado em 06.08.2026)

Contudo, a utilização de modelos avançados de inteligência artificial - designadamente redes neuronais profundas e aprendizagem automática não supervisionada (*machine learning*) - introduz a patologia da *opacidade algorítmica* ou o fenómeno da *black box* (caixa negra). Nestes sistemas, nem o próprio utilizador humano nem os programadores conseguem reconstituir com precisão os passos intermédios que levaram o modelo a produzir determinado resultado em face do *input* fornecido. A invocação genérica de uma fundamentação superficial e preponderantemente opaca não satisfaz o requisito constitucional da fundamentação expressa, traduzindo-se numa falta absoluta de motivação válida que fere o ato de invalidade.

Assim, é neste sentido, que o dever de fundamentação deve ser reforçado, no sentido de resultar aquilo que designamos por **"***fundamentação algorítmica reforçada*", ou seja, uma decisão que exige a verificação concomitante de duas dimensões de explicabilidade: uma *explicabilidade interna (ou material),* a qual é resultado de uma conceção clássica do Direito Administrativo, ou seja, exige-se a indicação clara e percetível das razões fácticas e legais que motivaram a decisão no caso concreto do administrado, demonstrando a relação de causalidade entre os dados do cidadão e o resultado final; e a *explicabilidade externa (ou funcional),* representando esta o vetor de inovação tecnológica, e pela qual se exige à Administração a obrigação de dar a conhecer a arquitetura do sistema, o código-fonte, as variáveis consideradas, os pesos e parâmetros de ponderação atribuídos e o histórico verificável da decisão (*logs*)[20] [21].

Apenas desta forma, será possível concretizar o desiderato de publicidade da atuação administrativa e a manutenção do valor constitucional da devida fundamentação dos atos administrativos, dando a conhecer ao seu destinatário o teor da decisão e as motivações factuais, jurídicas e estatísticas que levaram ao resultado exteriorizado[22].

---

20 Esta exigência de explicabilidade externa acolhe o ensinamento fixado pela jurisprudência europeia – designadamente pela Sentença do Supremo Tribunal Espanhol (Acórdão n.º 1119/2025, relativo à aplicação BOSCO) e a Sentença do Tribunal de Haia no caso *SyRI* (2020) -, as quais assentaram que a opacidade dos programas estatais sob o pretexto de propriedade intelectual ou segurança nacional viola o direito fundamental à transparência e à tutela jurisdicional efetiva.

21 Veja-se, neste sentido, quadro legal francês, designadamente o Artigo L. 311-3-1 do *Code des relations entre le public et l'administration*, o qual consagra um autêntico direito do cidadão a conhecer as regras que definem o tratamento algorítmico que o afeta.

22 Neste sentido, veja-se o nosso, LUÍS MANUEL PICA, *A Inteligência Artificial no Direito Tributários: Fundamentos e Limites Constitucionais*, Coimbra, Almedina, 2023, pp.181 e ss.; ANGELO GIUSEPPE OROFINO, "*The Implementation of the Transparency Principle in the Development of Electronic Administration*", *in*

O princípio da participação e a reconfiguração da audiência prévia digital

O princípio da participação, consagrado no artigo 267.º, n.º 5 da Constituição da República Portuguesa, e densificado no artigo 12.º do Código do Procedimento Administrativo, constitui a expressão da democracia procedimental. A sua principal concretização reside, entre outros, no direito de audiência prévia (artigo 121.º do Código do Procedimento Administrativo), o qual garante que o administrado tem o direito de se pronunciar antes de ser proferida uma decisão final suscetível de lesar os seus direitos ou interesses. A audiência prévia não constitui um mero rito formal ou uma cortesia burocrática, mas sim uma verdadeira exigência de *racionalidade comunicativa*. Serve para que o administrado traga ao procedimento factos atípicos, esclareça equívocos e influa materialmente na formação do juízo final da Administração.

A introdução desregulada de sistemas de inteligência artificial gera a transição de um procedimento *dialógico* para um procedimento *determinístico*. Ao operar com base em parâmetros padronizados e automatizados, a Administração corre o risco de suprimir tacitamente a audiência prévia, emitindo decisões instantâneas e remetendo o cidadão para reações litigantes *ex post*. A mera busca pela celeridade procedimental ou pela eficiência administrativa não encontra abrigo no elenco taxativo de exceções do artigo 124.º do Código do Procedimento Administrativo para dispensar a audiência dos interessados. A supressão não justificada do contraditório inquina o procedimento de vício de forma por preterição de formalidade essencial.

Neste sentido, não se advoga a rejeição do progresso tecnológico na Administração Pública, mas sim a reconfiguração funcional das garantias procedimentais. Por isso, propõe-se a (re)modelação de um *contraditório digital de participação*, assente nos seguintes pressupostos dogmáticos e procedimentais:

Fase da Audiência Digital	Mecanismo de Garantia Legal e Técnica
1. Notificação Algorítmica Prévia	O administrado é notificado eletronicamente da intenção de decisão automatizada, acompanhada da visualização clara dos dados processados e dos critérios algorítmicos aplicados.
2. Introdução de factos e dados novos	Concede-se uma plataforma para que o cidadão possa corrigir incorreções nos dados, introduzir ou trazer novos factos atípicos, ou invocar circunstâncias particulares não detetadas pela fórmula padrão.
3. Suspensão e Intervenção Humana	A submissão de alegações na audiência prévia suspende automaticamente o circuito automatizado, exigindo a intervenção obrigatória de um decisor humano (*human-in-the-loop*).

Veja-se que tal solução não é inovadora no sistema jurídico português, encontrando algumas semelhanças com o regime previsto no artigo 58.º-A do Código do IRS.

Nesta situação, a *notificação algorítmica* prévia materializa-se quando a Autoridade Tributária disponibiliza ao contribuinte, no Portal das Finanças, a sua declaração provisória de rendimentos acompanhada da respetiva demonstração de liquidação. Nesse instante, o algoritmo da máquina já cruzou milhões de dados provenientes de terceiros - desde entidades patronais a instituições bancárias -, mas a decisão final não se impõe de imediato. O contribuinte pode visualizar, com detalhe, o processamento de dados e os critérios aplicados aos seus rendimentos e deduções; num segundo momento, o *direito à introdução de factos novos* surge como a garantia de que o contribuinte pode alegar factos ou indicar desconformidades com os dados pré-preenchidos pela administração, caso estes não reflitam a complexidade da sua vida pessoal, como a alteração do agregado familiar, despesas de saúde não registadas ou factos erroneamente enquadrados; por fim, o mecanismo de *suspensão e intervenção humana* funciona como o garante do princípio do *human-in-the-loop*, pois no regime do IRS Automático, se o sujeito passivo não aceitar a proposta ou decidir apresentar a própria declaração com os dados inseridos pelo mesmo, a conversão automática da declaração e a prática automática do ato de liquidação fica sem efeito.

Ora, atendendo ao *princípio do controlo humano* estatuído no artigo 14.º do Regulamento (UE) 2024/1689 (*AI Act*), parece-nos uma solução capaz de assegurar que o algoritmo serve como ferramenta de apoio à decisão, mas nunca como substituto ilegítimo da ponderação jurídica humana. Isto é, a decisão acabará por ser totalmente imputada ao órgão decisor representado pelo agente do mesmo, e, portanto, capaz de alterar o teor da decisão a qualquer momento.

Garantias de imparcialidade, igualdade e proporcionalidade

O princípio da imparcialidade (artigo 9.º do Código do Procedimento Administrativo), em direta conexão com o princípio constitucional da igualdade (artigo 13.º da Constituição da República Portuguesa), exige que a Administração atue com isenção, objetividade e equidade na prossecução do interesse público.

Todavia, numa administração baseada em sistemas de inteligência artificial não pode sustentar a sua atuação na ideia de "*imparcialidade matemática*" ou na "*objetividade neutra*". Os sistemas de inteligência artificial são suscetíveis de incorporar e amplificar *enviesamentos discriminatórios* **(***algorithmic bias***)**, provenientes de dados de treino historicamente viciados, incompletos ou não representativos, ou na seleção desproporcional de variáveis e parametrizações matemáticas tendenciosas, ou, ainda, em modelos de aprendizagem que replicam e perpetuam desigualdades socioeconómicas, de género, raciais ou territoriais sob o manto de uma aparente neutralidade técnica.

Acresce a esta patologia o fenómeno altamente complexo do *bias propagation (ou amplificação de viés máquina-a-máquina***)**. Ocorre quando diferentes sistemas automatizados da Administração pública interagem entre si — por exemplo, quando um algoritmo de análise de risco fiscal importa automaticamente pontuações de risco geradas por um algoritmo de combate à fraude na segurança social. Se os modelos de origem contiverem distorções iniciais, a interoperabilidade sistémica multiplica em cadeia o viés discriminatório, criando uma teia de penalizações automatizadas onde se perde a rastreabilidade da injustiça originária.

Diante desta realidade, o dever de imparcialidade deixa de se aferir apenas pela postura ética do agente público individual, passando a exigir da Administração uma garantia de *neutralidade sistémica da tecnologia*. Para tanto, torna-se obrigatória a realização

continuada de auditorias de equidade, testes de representatividade de dados e avaliações de impacto que permitam assegurar a plena tutela dos direitos fundamentais dos administrados, em conformidade com os artigos 9.º e 10.º do Regulamento (UE) 2024/1689.

Por outro lado, não se poderá esquecer que a integração de sistemas de inteligência artificial no seio da Administração Pública deve ser feita em respeito estrito pelo princípio da proporcionalidade. Conforme consagra o artigo 7.º do Código do Procedimento Administrativo, o princípio da proporcionalidade opera como limite imanente ao exercício de qualquer poder público, pois numa administração algorítmica, a adoção de soluções tecnológicas não pode deixar de ser integrada com pleno respeito pela adequação, necessidade e proporcionalidade *stricto sensu.*

Ou seja, o sistema de inteligência artificial adotado deve revelar-se idóneo e juridicamente adequado para prosseguir o fim público específico fixado na lei (*adequação),* bem como a Administração tem o ónus de demonstrar que os procedimentos adotados (*v.g.* modelos preditivos ou de criação de perfis) são os meios menos intrusivos dos direitos fundamentais dos sujeitos para alcançar aquele objetivo (*necessidade),* exigindo-se um juízo de ponderação justa entre os ganhos de eficiência administrativa obtidos pela automação e o eventual sacrifício imposto às posições jurídicas dos administrados (*v.g.* privacidade, não discriminação e contraditório).

Por isso, o desrespeito por qualquer uma destes direitos fundamentais configura uma verdadeira disrupção que não pode ser admitida num Estado de Direito, correspondendo este ao dever pela preservação da ordem jurídica, dos direitos fundamentais e pelo respeito pelas posições jurídicas adquiridas pelos administrados. Uma Administração Pública que, em prol da eficiência, consagre verdadeiros instrumentos de atuação lesivos dos direitos dos administrados, não pode deixar de corresponder a uma verdadeira invalidade do procedimento e do ato resultante deste.

V. CONCLUSÃO

A expansão da Inteligência Artificial no seio da Administração Pública representa um ponto de viragem histórico que não deve culminar numa rejeição da inovação tecnológica, nem tão-pouco num deslumbramento cego pela eficiência computacional.

O verdadeiro desafio do Direito Administrativo contemporâneo reside na capacidade de evoluir a sua dogmática e renovar os seus institutos sem abdicar, em momento algum, das conquistas civilizacionais que estruturam o Estado de Direito material.

É na interseção entre a preservação dessas garantias e a irrupção das novas realidades tecnológicas que se desenha a necessária reconfiguração normativa e doutrinária do procedimento. Desde logo, impõe-se a afirmação dogmática do princípio da reserva de decisão humana, consagrando legislativamente a proibição absoluta de atos totalmente automatizados sempre que a atuação administrativa envolva o exercício de poder discricionário, a valoração de conceitos indeterminados ou a apreciação casuística da equidade, em moldes semelhantes aos trilhados pela Lei de Procedimento Administrativo alemã.

A esta limitação substancial acresce a imperiosa consagração de um dever de fundamentação algorítmica reforçada, o qual deve operar perante uma dupla dimensão, exigindo a explicabilidade interna, traduzida na motivação percetível das razões fácticas e jurídicas aplicadas ao caso concreto, e a explicabilidade externa, fundada na transparência do próprio sistema, dos seus parâmetros e do seu código-fonte, em plena sintonia com os ditames europeus.

Por outro lado, a preservação da dimensão democrática do procedimento exige a instituição de um genuíno contraditório digital. Trata-se de reconfigurar o tradicional direito de audiência prévia para garantir que o administrado é atempadamente notificado dos critérios algorítmicos aplicáveis, assegurando que a invocação de exceções ou a junção de novos elementos determine a suspensão do automatismo e obrigue a uma reavaliação por um decisor humano (*human-in-the-loop*).

Esta nova dinâmica assenta, reflexivamente, na afirmação da tese da autoria institucional e na garantia de uma neutralidade sistémica, reconhecendo que a responsabilidade jurídica e política pelos atos produzidos recai invariavelmente sobre a instituição pública titular, à qual

cumpre realizar avaliações *ex-ante* de impacto sobre os direitos fundamentais e promover auditorias periódicas destinadas a erradicar o viés algorítmico e os seus fenómenos de amplificação em cadeia.

Por fim, este edifício analítico completa-se com a densificação doutrinária e jurisprudencial dos clássicos vícios de validade do ato administrativo, retraduzidos agora sob as formas de incompetência tecnológica, vício de forma por opacidade, violação de lei por desvio de parametrização e desvio de poder algorítmico. Em última análise, apenas mediante esta submissão estrita dos instrumentos tecnológicos ao bloco da legalidade e aos princípios procedimentais será possível assegurar que a modernização da Administração Pública serve a justiça material, a transparência e a salvaguarda indisponível dos direitos dos cidadãos.



O RISCO OCULTO DA INTELIGÊNCIA ARTIFICIAL: COMO A AUTOMATIZAÇÃO PODE DEGRADAR A CAPACIDADE COGNITIVA, LABORAL, ORGANIZACIONAL E SECURITÁRIA

FERNANDO CAVALEIRO ÂNGELO *

* Mestre em Estudos de Informações e Segurança (Universidade de Salford, Reino Unido)

RESUMO:

O presente ensaio analisa a transição da Inteligência Artificial (IA) de uma fase meramente experimental para uma integração sistémica na estrutura económica e social contemporânea. Através de uma revisão das transformações no mercado de trabalho, nos modelos educativos e nas dinâmicas de segurança nacional e defesa, examina-se a natureza ambivalente desta tecnologia. Se, por um lado, a IA potencia ganhos de produtividade e capacidades operacionais sem precedentes, por outro, introduz riscos de desigualdade sócio-económica, desinformação e a emergência de uma "lassidão cognitiva". O trabalho conclui que a mitigação destes riscos exige um investimento estratégico na literacia digital e o reforço do pensamento crítico, garantindo que a tecnologia atue como um complemento, e não um substituto do tradicional processo cognitivo humano.

Palavras-chave: Inteligência Artificial; Automatização; Capacidade Cognitiva; Mercado de Trabalho; Educação; Segurança e Defesa; Drones; Desinformação; Literacia Digital; Governação Tecnológica; FIMI.

ABSTRACT:

This paper analyses the transition of Artificial Intelligence (AI) from a merely experimental phase to a systemic integration within contemporary economic and social structures. Through a review of transformations in the labour market, educational models, and national security and defence dynamics, it examines the ambivalent nature of this technology. Whilst AI yields unprecedented productivity gains and operational capabilities on the one hand, it introduces risks of socio-economic inequality, disinformation, and the emergence of "cognitive lassitude" on the other. The paper concludes that mitigating these risks requires strategic investment in digital literacy and the strengthening of critical thinking, ensuring that technology acts as a complement to, rather than a substitute for, traditional human cognitive processes.

Keywords: Artificial intelligence; Automation; Cognitive capacity; Labour market; Education; Security and defence; Drones; Disinformation; Digital literacy; Technological governance; FIMI.

I. INTRODUÇÃO

A Inteligência Artificial (IA) consolidou-se, na sociedade contemporânea, como um pilar estruturante que transcende o espetro meramente técnico para influenciar as dimensões profissionais, culturais e sociais. Deixando de ser um domínio de cariz experimental, a IA passou a sustentar o funcionamento de setores críticos, como a saúde, a justiça e a defesa, impulsionada por investimentos globais massivos que, em 2024, atingiram níveis recorde. Esta integração sistémica promete uma otimização de processos e ganhos de eficiência que outrora levariam meses a concretizar, mas traz consigo uma natureza ambivalente.

O paradigma atual não se limita a uma evolução tecnológica, mas sim a uma profunda reorganização económica e social. Embora a automação possa gerar novos postos de trabalho, subsistem receios legítimos quanto ao "colapso do emprego" e ao agravamento das disparidades entre detentores de capital e trabalhadores. Paralelamente, a dependência crescente de sistemas automatizados expõe as instituições a fenómenos de desinformação e manipulação, colocando em causa a soberania do pensamento crítico. O objetivo deste ensaio é, portanto, discutir como a adaptação das instituições e dos enquadramentos regulatórios se torna fundamental para maximizar os benefícios desta "era da IA" sem comprometer a estabilidade social e a integridade cognitiva dos indivíduos.

II. NOVO PARADIGMA

De acordo com o relatório de IA de 2025 publicado pela Universidade de Stanford, na Califórnia, o ano de 2024 consolidou a transição da IA de uma fase experimental para uma integração sistémica na economia global. O investimento privado global em inteligência artificial registou um crescimento expressivo de 44,5%, atingindo 252,3 mil milhões de dólares, com a IA generativa a representar mais de 20% deste total. Paralelamente, a adoção organizacional disparou de 55% para 78%, enquanto o uso de IA generativa duplicou para 71% das empresas. No mercado laboral, a tecnologia alegadamente impulsionou ganhos de produtividade mensuráveis (entre 10% e 45%) e demonstrou um efeito equalizador, beneficiando proporcionalmente trabalhadores menos qualificados. Em termos globais, os Estados Unidos da América mantêm a hegemonia no investimento privado, alegadamente

superando a China. Contudo, Pequim domina a robótica industrial, sendo responsável por 51,1% das instalações mundiais, segundo o mesmo relatório. A substituição de trabalhadores por máquinas, robôs ou aplicações de IA pode induzir a extinção de certos nichos do mercado laboral. Com o passar dos tempos, pode haver o nascimento de algumas novas profissões, ou o aumento de outras atividades profissionais.

A par com estes desenvolvimentos, com os consequentes ajustes em diversas atividades laborais, a facilidade e rapidez com que a tecnologia IA dá respostas, pode induzir o ser humano a acomodar-se e não efetuar uma análise crítica daquilo que é apresentado. Isto pode, em resultado, tornar aqueles que não querem fazer um esforço analítico em meros ignorantes e preguiçosos cognitivos. Nunca se deve descartar as frequentes ondas de desinformação ou manipulação que a IA pode induzir nas massas, quando esta está a ser operada por atores estatais ou não-estatais, para levar a cabo agendas que operacionalizam os seus intentos. Se não existir um escrutínio e análise crítica das imagens, gravações ou textos que são massivamente difundidos na internet, com amplo uso das redes sociais para atingir populações alvo, as pessoas ficam vulneráveis a campanhas de desinformação e propaganda produzidas a partir de ferramentas de IA. E se esta população perde a confiança nas entidades estatais ou em instituições até então credíveis, poder-se-á inferir que o caminho para a instabilidade social é um cenário bastante provável.

III. MILAGRE ARTIFICIAL

Um dos termos mais escutados e propalados nos dias de hoje é a inteligência artificial. É a solução para tudo, começando a constituir-se como o novo remédio para todos os males. Não há uma intervenção, palestra, apresentação, lição ou outro qualquer evento que não contenha algures uma menção à IA. Até no domínio militar começa a sobressair o milagre da inteligência artificial para aniquilar os opositores e capturar os malfeitores, ou gerar soldados de silício para colmatar as dificuldades de recrutamento de seres humanos. Em todos os ramos de atividade, excetuando aquelas onde o ser humano ainda é o elemento central, como canalizador, eletricista, pedreiro, pescador, bombeiro, médico, enfermeiro, entre outras tantas não menos importantes, este novo milagre civilizacional que atualmente

movimenta triliões de euros por este mundo fora, pode estar a empurrar para o desemprego milhões de pessoas em todo o mundo.

Em muitas situações, o ser humano constituiu-se como um elemento descartável, a troco da alegada eficiência e eficácia. Muitas das vezes é puramente uma questão de redução de custos e alavancagem de lucros. O ser humano é uma máquina muito mais complexa e dispendiosa do que um simples ser virtual e não tangível. Mas se começarmos a descartar os seres humanos, vamos metê-los a fazer o quê? Se os empurrarmos para o desemprego, as máquinas e robôs que corporizam a inteligência artificial e que não descontam para a segurança social, como é que os Estados vão financiar os desempregados?

Obviamente que as extrapolações são sempre um exercício perigoso e podem percecionar conceitos errados e fora de contexto. Existirão, atualmente, trabalhos para identificar a alteração da dinâmica no mercado laboral, decorrentes da extinção de alguns setores e o aparecimento de novos que advêm da nova revolução do século XXI. Aqui poder-se-ia fazer um paralelo com a revolução industrial no século XIX, onde naquele tempo a passagem de métodos de produção manual para processos mecânicos também trouxe muita apreensão e transferência de competências e, em alguns casos, desemprego. Tal como o mundo se soube adaptar naquela época, também agora vai ter de haver um processo transformacional para uma nova realidade.

IV. TRANSFORMAÇÃO LABORAL

O panorama global do mercado de trabalho até 2030 será definido por uma transformação profunda impulsionada pela IA, conforme indicam análises recentes. Um relatório da *Goldman Sachs* de 2025 estima que a IA poderá automatizar cerca de 25% de todas as horas de trabalho nos Estados Unidos da América e na Europa, expondo dois terços das ocupações atuais a algum grau de automação e afetando potencialmente 300 milhões de funções a tempo inteiro a nível mundial.

Esta perspetiva é complementada e detalhada pelo Relatório sobre o Futuro dos Empregos 2025 do Fórum Económico Mundial (WEF), que prevê uma rotação estrutural do mercado de trabalho de 22%, resultando na criação de 170 milhões de novos postos de trabalho (14%) e no deslocamento de 92 milhões de funções existentes (8%). Apesar das preocupações com

a perda de postos, projeta-se um crescimento líquido positivo de 7%, ou 78 milhões de empregos, impulsionado por ganhos de produtividade que a *Goldman Sachs* estima poderem atingir os 15% nos mercados desenvolvidos, adicionando 7 biliões de dólares ao PIB global na próxima década.

As funções de apoio administrativo, jurídico e financeiro são identificadas como as mais vulneráveis à automação, com o relatório do WEF a confirmar que cargos de secretariado, caixas de banco e supermercados, e técnicos de introdução e análise de dados, enfrentarão os declínios mais acentuados em números absolutos. Simultaneamente, observa-se uma mudança significativa na fronteira entre humanos e máquinas: prevê-se que a percentagem de tarefas executadas exclusivamente por pessoas caia de 47% em 2025 para 34% até 2030. No entanto, o foco está a afastar-se da substituição direta para a chamada "aumentação", onde a tecnologia complementa e expande as capacidades humanas em vez de as eliminar totalmente. Setores como a saúde e o setor público deverão liderar esta tendência de colaboração homem-máquina, enquanto a automação mais profunda será prevalecente em áreas como os serviços financeiros e eletrónicos.

Esta evolução tecnológica exige uma adaptação sem precedentes nas competências dos trabalhadores, estimando-se que 39% das competências centrais atuais sofram transformações ou se tornem obsoletas até ao final da década. O pensamento analítico permanece a competência mais essencial, procurada por sete em cada dez empresas, mas as ligadas à IA, *Big Data* e cibersegurança são as que apresentam o crescimento mais acelerado na procura. Algumas profissões essenciais na economia da prestação de cuidados e na educação, como médicos, enfermeiros e professores, não são afetadas. O WEF destaca que, embora a IA generativa tenha um elevado potencial para substituir tarefas baseadas em conhecimento teórico e manipulação digital (como escrita e matemática), continua a ter limitações em funções que exigem empatia humana, julgamento complexo ou execução física manual. Perante este cenário, as lacunas de competências são apontadas por 63% dos empregadores como a principal barreira à transformação empresarial, o que leva 85% das organizações a planearem priorizar a requalificação da sua força de trabalho.

Para navegar estas transições, as empresas estão a adotar estratégias que ultrapassam a mera remuneração, focando-se no apoio à saúde e ao bem-estar dos funcionários como a estratégia prioritária para atrair talento, à frente do aumento salarial ou da progressão interna. Além

disso, observa-se um compromisso crescente com iniciativas de diversidade, equidade e inclusão, com 83% das empresas a reportar programas ativos, um aumento significativo face aos 67% registados em 2023, sendo as mulheres o grupo prioritário nestas medidas. O relatório do WEF conclui que o futuro do trabalho até 2030 sugere um equilíbrio onde a IA atua como um motor de produtividade e criação de emprego, exigindo um investimento massivo em capital humano através de uma abordagem centrada nas competências para garantir uma prosperidade partilhada.

Em termos nacionais, num estudo desenvolvido pela Fundação Francisco Manuel dos Santos em parceria com o Instituto Superior Técnico, aponta para os setores da educação, como professores e educadores, e finanças e contabilidade a serem setores com potencial de benefício e/ou ascensão. Quanto aos que estão em risco, ou em alguns casos mesmo colapso, são salientados os setores da logística, agricultura e serviços de apoio ao cliente, como o caso dos operadores de *call center*. Alguns estudos são unânimes a prever que cerca de 30% dos empregos em Portugal estão em risco de extinção ou transformação significativa devido à IA. Setores com tarefas repetitivas, como vendas e operação de máquinas, entre outros, são os mais vulneráveis. Este processo de automação está necessariamente a criar novas oportunidades, como gestores de redes sociais, especialistas em dados e técnicos de IA. A requalificação profissional é, desta forma, apontada como fundamental para a transição dos trabalhadores de áreas em potencial extinção para outras em ascensão. Existe, desde logo, a necessidade de uma adaptação ao nível educativo e formativo. A questão que se impõe neste momento é se as escolas e universidades portuguesas estão a adaptar o seu portfólio curricular a estas novas tendências?

V. EFEITOS NA EDUCAÇÃO E TREINO

Reserva-se algumas dúvidas sobre estas necessárias linhas de ação, pois ao consultar a oferta de cursos oferecidos nos respetivos catálogos, constata-se que algumas licenciaturas muito em breve não mais vão fazer do que formar jovens para o mercado de desemprego. Alguns cursos não vão ter saída profissional, pelo que urge uma transformação curricular para fazer face às novas tendências impostas pela IA. Outra das ideias que surgiu de imediato foi colocar avatares a darem ações de formação, medida esta que pouparia uma grande

quantidade de professores. Professores quase reais, com aparências imaculadas, discursos perfeitos e apoios gráficos que apresentam os conteúdos programáticos com as técnicas formativas sem qualquer falha, são uma resposta aos mais dispendiosos professores de carne e osso. Se existir alguma alteração nos programas, basta alterar o conteúdo e numa fração de segundos temos os avatares a lecionar a matéria atualizada. Mas este mundo perfeito não cabe no sistema educativo, onde a presença física do professor humano é requerida, pois em matéria de ação formativa, o contacto visual, expressões, sentimentos, maneirismos controlados, e capacidade de adaptação à atmosfera em sala de aulas, ainda prevalece sobre os avatares. Isto não quer dizer que os avatares não sejam úteis no processo educativo, onde para algumas ações de apoio ao esforço letivo do professor, eles podem assumir uma função de complementaridade.

O centro do processo educativo são os alunos. E aqui levanta-se um dos maiores desafios. Como estimular os alunos a não se transformarem em meros ignorantes altamente dependentes do conhecimento que só a inteligência artificial disponível nos seus telemóveis ou computadores lhes pode facultar? Ao serem estimulados desde tenra idade a estarem dependentes das ferramentas de inteligência artificial para formular questões, fazerem pesquisas e análises, e a traçar conclusões que são geradas em toda a sua extensão por computadores, tornam-nos uns meros preguiçosos mentais sem qualquer capacidade crítica para avaliar aquilo que a máquina lhes proporciona em poucos segundos. Em diversas situações o ser humano tende a acomodar-se às circunstâncias. O cérebro humano é a máquina mais complexa que existe. Por norma, as ligações que o cérebro faz tendem sempre a seguir um padrão que assenta num processo cognitivo construído através da sua génese social, educacional e organizacional. Quanto mais robustos forem estes alicerces, maior é o espírito crítico que o indivíduo tende a ter para decifrar, desconstruir e confirmar a informação que a inteligência artificial lhe proporciona. Caso contrário, aplica-se a norma anglo-saxónica de "*garbage in, garbage out*!"

Mas tudo isto não quer necessariamente dizer que a IA não seja uma ferramenta importante para auxiliar os estudantes no seu ciclo formativo. Claramente que as horas passadas na biblioteca a fazer pesquisa, procurando construir o conhecimento a partir do folhear de centenas de livros, artigos científicos, arquivos e demais informação, estão resumidas a uma simples interação computacional. Aquelas horas são agora transformadas em poucos

segundos. A lentidão de colocar cada tijolo na parede do conhecimento agora é substituída por uma parede pré-fabricada que só requer uma validação e aperfeiçoamento para se tornar robusta. Se a utilização dessa informação não for questionada com um espírito crítico, confrontando a credibilidade das fontes, cruzando com outra informação, atestando a sua validade científica, a ignorância e lassidão vão lentamente tomando o controlo dos seres humanos. O ser humano tem de combater a tentação de ser um mero executor e não um ser pensante, que em última análise todo o processo termina com a sua aceitação, ou rejeição, do que lhe está a ser proposto ou apresentado.

VI. APLICABILIDADE NA SEGURANÇA NACIONAL

Em matéria de segurança nacional, mormente nos pilares da segurança e defesa, a IA adquire cada vez mais uma importância ao nível dos agentes computacionais e de robótica que muito coadjuvam a missão de combater as ameaças e riscos à soberania nacional. Muitos dos sistemas que agora substituem homens e mulheres são mais eficazes e eficientes, proporcionando um conhecimento situacional em tempo real. Se há desvantagens em matéria de sustentabilidade de alguns destes sistemas, tal como se observou durante o apagão energético que assolou Portugal em 2025, as vantagens ainda continuam a ser maiores do que empregar pessoas para executar todas as tarefas que estes podem proporcionar. A vulnerabilidade de muitos destes sistemas aos ciberataques é o calcanhar de Aquiles de alguns deles. Por vezes a escassez de orçamento para executar as necessárias manutenções e atualizações sob um ponto de vista da cibersegurança acaba por degradar os escudos de proteção que se querem imunes a tentativas de intrusão e disrupção. Os seres humanos também estão dotados de muitas vulnerabilidades, tais como a doença, incompetência e negligência. E contra isso não há muito a fazer, senão cuidados de saúde, formação e treino e, em último caso, o despedimento.

Os sistemas de comando e controlo das unidades militares estão cada vez mais a desenvolverem-se segundo modelos comerciais *off-the-shelf*, com IA a assumir algoritmos de interpretação e interligação que faz com que milhões de dados, de diferentes proveniências e múltiplos formatos, sejam transformados em informação gráfica digerível ao decisor. A matriz taxonómica do militar requer sempre redundância quando se trata de comando e

controlo. Os vários escalões de comando, a sua interligação e coordenação, assim o requer. Quando surgiram as cartas náuticas eletrónicas, vulgo *ecodis*, o sistema para além de ter a cartografia atualizada, permitia fazer de forma intuitiva e fácil a navegação sem recurso a qualquer carta náutica de papel, compasso de pontas, esquadros de navegação, lápis e borracha. Mas a decisão foi manter o processo manual, pois seria a forma redundante caso este colapsasse. E esta decisão foi muito acertada. Não há muitos anos atrás, um navio de guerra teve um acidente fatal devido a uma acumulação de erros. Na génese destes erros esteve a confiança total no sistema eletrónico de navegação. Não houve um espírito crítico do oficial de quarto à ponte para perceber os erros que o sistema lhe estava a proporcionar, não fazendo juízo dos seus sensores naturais (visão) e das posições marcadas manualmente na carta náutica que deveria ter feito. Mas esta redundância tem custos. Quer em termos materiais, quer em tempo gasto com os recursos humanos em ações de formação e treino. No passado muito recente, com a falta de orçamento para colocar navios no mar a navegar, optou-se por formar as equipas de navegação em simuladores que replicavam todos os sistemas a bordo. Esta realidade virtual é importante para proporcionar entrosamentos de equipas de ponte dos navios, em termos de procedimentos e ações de comando e controlo, mas não consegue substituir as perícias adquiridas num domínio real. Tal como a presença do professor na sala de aulas. É necessário sentir as condicionantes e sensações reais num determinado momento e espaço físico, percebendo as reações causa-efeito das ordens dadas perante as imposições atmosféricas e hidrográficas. Estas duas opções não se substituem. Pelo contrário, complementam-se.

Esta complementaridade tem sempre o ser humano no centro do processo. A rapidez de processamento que é necessária numa era em que a velocidade, quantidade e diversidade dos dados aumenta de forma exponencial, exige sistemas que ajudem o ser humano a tomar a decisão final. A guerra moderna está a ser transformada radicalmente pela utilização da inteligência artificial, servindo o conflito na Ucrânia e no Irão como um "laboratório de guerra de IA" onde todos os envolvidos utilizam estas tecnologias na recolha de informações e nas operações cibernéticas, cinéticas e de sistemas não tripulados. Estes teatros vão incrementar ainda mais a corrida armamentista global que se tem observado no passado recente, liderada pelos Estados Unidos e pela China, com a Rússia a investir fortemente para colmatar a lacuna tecnológica.

Enquanto os EUA integram a IA através de iniciativas como o *Project Maven* e o *Chief Digital and Artificial Intelligence Office* (CDAO) para manter a "superioridade de decisão", a China vê a IA como o pilar central para tornar as suas forças armadas de "classe mundial" até meados do século. A Rússia, por sua vez, prioriza a IA como um multiplicador de força, com o objetivo de automatizar 30% do seu equipamento militar. Um aspeto crucial desta transformação é a inversão do fluxo de inovação: se anteriormente as tecnologias militares derivavam para o domínio civil, hoje a inovação em IA é predominantemente impulsionada pelo setor privado.

Empresas tecnológicas globais como a Alphabet, Microsoft, Palantir e a chinesa Alibaba tornaram-se atores geopolíticos cujas capacidades e recursos podem alterar o curso de conflitos, operando por vezes com uma autonomia sem precedentes. Um exemplo notório é a utilização dos satélites Starlink, da SpaceX, para fornecer comunicações críticas durante a guerra na Ucrânia. No contexto europeu, a União Europeia (UE) tem procurado fortalecer a sua "*autonomia estratégica*" e "*soberania tecnológica*" perante estas rivalidades geopolíticas. Através de diversos mecanismos e fontes de financiamento, a UE dedica uma percentagem do seu orçamento a tecnologias disruptivas, incluindo a IA em áreas como a defesa cibernética e sistemas autónomos cooperativos.

A integração da IA na defesa levanta dilemas éticos profundos, nomeadamente sobre a prestação de contas, a conformidade com o direito internacional humanitário e o risco de escalada de conflitos devido à redução da supervisão humana. Aqui o debate sobre a regulação de Sistemas de Armas Letais Autónomas (*LAWS*) assume uma posição central. Enquanto países como os EUA e a Rússia se opõem a uma proibição preventiva, o Parlamento Europeu tem apelado consistentemente à proibição de armas que operem sem "*controlo humano significativo*".

Embora o Regulamento da Inteligência Artificial (*AI Act*) da UE seja um marco regulatório mundial, este exclui explicitamente as aplicações militares. No entanto, a sua abordagem baseada no risco poderá influenciar futuras normas de defesa e o desenvolvimento de tecnologias de duplo uso. Em última análise, a posição do Parlamento Europeu reconhece a importância estratégica da IA, mas sublinha a necessidade de diretrizes éticas robustas e de uma governação internacional que impeça o uso irresponsável destas tecnologias. A máquina

é imputável, pois em caso de acidente não se pode acusar em tribunal o sistema de ter atuado de forma negligente ou criminosa.

Com tanta automação, com tantos processos que são efetuados sem controlo humano, mesmo que o algoritmo tenha sido desenhado segundo um critério estabelecido por seres humanos, o resultado pode ser catastrófico se o decisor não entender toda a dinâmica que está por detrás do fluxograma que orienta um determinado sistema.

VII. MUDANÇA NAS OPERAÇÕES MILITARES

Com os cenários de guerra da Ucrânia, depois da invasão da Federação Russa, os modelos de estudo da guerra ensinados nas academias têm sido colocados à prova. As manobras e dispositivos táticos de campanha, as cadeias de comando e controlo, os sistemas ofensivos e defensivos nos diversos domínios (aéreo, terrestre, marítimo, espacial e cibernético), a utilização de guerra eletrónica, e a utilização e gestão de mísseis de curto-médio-longo alcance, entre tantos outros, tornaram-se ineficazes perante sistemas comerciais de baixo custo e de fácil aquisição e sustentação. Está-se a falar do atual conceito de utilização de drones, que alteraram sobremaneira o paradigma dos sistemas de muitos milhões de euros e a táticas de manobra militares que até então se pensava amplamente testadas e comprovadas.

O ambiente de segurança contemporâneo está a sofrer uma transformação profunda, impulsionada pela industrialização em larga escala e pela integração dos Sistemas Aéreos Não Tripuladas (UAS) na guerra russo-ucraniana. Os drones deixaram de ser ferramentas periféricas de reconhecimento para se tornarem também em instrumentos centrais da guerra híbrida, com um conjunto de ações hostis conduzidas em períodos de paz formal, recorrendo a tecnologias de duplo uso para desestabilizar ou coagir um adversário sem desencadear um conflito armado aberto. Na Ucrânia, esta mudança traduz-se numa verdadeira "*alteração de paradigma*", em que sistemas de baixo custo e difícil deteção conseguem sobrecarregar as tradicionais e dispendiosas arquiteturas de defesa aérea. O conflito transformou-se, assim, num laboratório em escala real para tecnologias não tripuladas, com uma ampla utilização de IA, passando de uma lógica centrada nas plataformas bélicas convencionais para operações

orientadas pela fusão de dados, onde os drones funcionam como nós dinâmicos numa malha integrada de deteção e ataque.

Um dos desenvolvimentos mais marcantes no teatro ucraniano é a ascensão dos drones *First Person View* (FPV), originalmente associados ao desporto de corridas de drones. Estes sistemas recorrem a *firmware* de código aberto e a componentes comerciais amplamente disponíveis, o que os torna suficientemente baratos para serem produzidos numa escala em massa. Em 2025, a produção diária destes drones na região ultrapassava já as 6.000 unidades. O seu impacto é disruptivo: permitem ataques de elevada precisão contra alvos de alto valor, desde veículos fortemente blindados até soldados individuais, a um custo bastante mais reduzido do despendido em munições convencionais.

O impacto operacional é profundo. Alguns relatórios publicados em 2026 indicam que cerca de 80% das baixas no campo de batalha são causadas por ataques com drones. Os drones FPV são normalmente operados por equipas de três a quatro pessoas e podem atingir distâncias superiores a 20 quilómetros quando apoiados por repetidores de sinal rádio para os controlar. Para contrariar a sua operação, têm sido empregues bloqueadores multibanda, sistemas de guerra eletrónica de alta potência e barreiras físicas, como redes ou estruturas metálicas sobre rotas logísticas principais. Contudo, surgiu uma nova geração de drones FPV com ligação por fibra ótica, completamente imunes à guerra eletrónica, uma vez que utilizam ligações físicas para controlo e transmissão de dados, não emitindo sinais de radiofrequência detetáveis por sistemas de inteligência de sinais (SIGINT). E é aqui que a IA está a assumir uma função central e decisiva no conflito.

Para além das capacidades de ataque direto, os drones tornaram o campo de batalha moderno praticamente "transparente", através de operações persistentes de pesquisa, vigilância e ataque. Os drones operacionais como o Supercam S350 e o Orlan-10 apresentam envergaduras significativas e autonomias de voo que podem atingir as oito horas, permitindo monitorizar áreas até aos 600 quilómetros de distância. Estas plataformas transportam frequentemente sistemas que permitem visão térmica, telemetria e designação laser para guiar munições de artilharia. Por exemplo, o Orlan-30 é frequentemente descrito como o "*melhor aliado da artilharia*", devido à sua capacidade de viabilizar fogos de elevada precisão.

Os drones comerciais, sobretudo de fabricantes como a DJI - que detém cerca de 94% da quota de mercado neste conflito - são também amplamente utilizados em missões de pesquisa, vigilância e ataque. No entanto, estes sistemas apresentam riscos significativos para os operadores, uma vez que protocolos como o OcuSync transmitem identificadores únicos não encriptados e coordenadas do ponto de origem, passíveis de serem intercetados por ferramentas como o Aeroscope. Para mitigar estes riscos, os utilizadores recorrem frequentemente a modificações de *hardware*, como a remoção de módulos GNSS[1] ou a instalação de *firmware* alternativo, de modo a ocultar a sua assinatura e operar em zonas restritas. Muitas destas operações de adaptação são efetuadas através do uso de IA.

A ameaça estende-se à profundidade estratégica do território através de drones de ataque de longo alcance, como o Geran-2 (Shahed-136). Estes sistemas podem transportar ogivas até 90 quilogramas de explosivos e atingir distâncias na ordem dos 2.500 quilómetros. No final da primavera de 2025, a Rússia produzia cerca de 170 drones Geran-2 por dia, apoiada por infraestruturas industriais de grande escala, como a fábrica de Yelabuga, que tem uma implantação territorial numa área de aproximadamente 10 quilómetros quadrados. Estes ataques visam esgotar os stocks de mísseis de defesa aérea, que são extremamente dispendiosos, e destruir infraestruturas críticas.

Um elemento central desta ameaça estratégica é o uso de engodos, como o UAV Gerbera. Trata-se de uma plataforma de baixo custo, estimado em menos de 10.000 dólares comparativamente aos 50.000 dólares de um Shahed, concebida para saturar os sistemas de defesa aérea, simulando aeronaves maiores no radar através de lentes de Luneburg. Estes engodos não possuem ogivas, o que reduz o risco de escalada militar imediata, mas obriga os defensores a revelar os seus procedimentos de resposta e a consumir recursos. São considerados uma ameaça híbrida significativa, uma vez que permitem negar plausivelmente a responsabilidade, podendo uma incursão ser atribuída a um erro de navegação.

À medida que a guerra eletrónica se torna mais sofisticada, os fabricantes de drones têm vindo a evoluir as suas tecnologias de navegação e conectividade para manter a eficácia

1 Global Navigation Sattelite System (GNSS) é o cérebro da navegação do drone, permitindo-lhe a localização exata em tempo real através dos sinais de várias constelações de satélites. Em vez do sistema de posicionamento GPS (Global Positioning System), os drones modernos utilizam módulos GNSS para captar sinais de múltiplos sistemas simultaneamente, o que aumenta a fiabilidade e evita a perda de sinal caso uma constelação de satélites tenha falhas.

operacional. As antenas do tipo *Controlled Reception Pattern Antennas* (CRPA) foram introduzidas para mitigar interferências e ataques de *spoofing*[2], atenuando sinais maliciosos e reforçando os dados legítimos provenientes de satélites. Em resposta, têm sido desenvolvidos sistemas distribuídos de contramedidas de guerra eletrónica, assentes em infraestruturas fixas com alimentação contínua e ligação à internet, capazes de executar operações de *spoofing* em larga escala.

A conectividade está também a evoluir no sentido das comunicações por satélite (SatCom) e do uso de modems celulares. Os sistemas SatCom de órbita baixa, como o Starlink, oferecem elevadas taxas de transmissão e são extremamente difíceis de neutralizar, enquanto os modems celulares permitem a transmissão em tempo real de dados de telemetria e reconhecimento através de infraestruturas civis existentes. Paralelamente, o desenvolvimento de redes em malha (*mesh networks*) possibilita que grupos coordenados de drones autónomos, os *vulgo* enxames, comuniquem e se adaptem sem controlo centralizado, transferindo a gestão da missão do operador humano para o sistema controlado por inteligência artificial. Todos estes desenvolvimentos contribuem para a sistematização de práticas para aumentar a sobrevivência dos drones através da extensa utilização da IA em operações autónomas.

Outro exemplo da mudança de paradigma deu-se em finais de setembro de 2025, quando uma equipa liderada por militares ucranianos que durante um exercício "afundou" um navio de guerra da NATO com recurso a pequenos drones. O jornal *Frankfurter Allgemeine* noticiou que uma fonte ucraniana que participou no exercício revelou que eles derrotaram consistentemente as forças da aliança em cinco cenários distintos, que incluíam a proteção de portos e a defesa e ataque de navios. Isto dá que pensar sobre a projeção de forças militares para um teatro de operações hostil, onde os seus sensores e armas de milhões de euros não conseguem detetar e abater uma ameaça que custa apenas alguns milhares de euros. Aqui a inteligência artificial joga um papel fundamental no desenho e construção de novos sensores e sistemas de armas que possam anular e/ou atenuar uma ameaça que também utiliza a lA para escapar às deteções e atacar sem ser abatido.

2 Falsificação ou mistificação.

VIII. CONTRIBUIÇÃO PARA A AMEAÇA HÍBRIDA

Do que no passado não muito distante se apelidava de guerra irregular, ou operações cobertas, os atuais desenvolvimentos geopolíticos e geoestratégicos fomentaram um conceito mais sofisticado e abrangente em todo o espetro da atividade político-social e militar: a guerra híbrida. A ameaça híbrida procura atingir objetivos estratégicos por intermédio da influência do processo de decisão, reduzindo ou atingindo o centro de gravidade do seu alvo, através de ações coordenadas e sincronizadas. Explora todas as ações para atingir as vulnerabilidades sistémicas do alvo, identificando os limiares de deteção e de atribuição, tal como as fronteiras existentes entre guerra e paz, interno e externo, e público e privado. Os vetores de ataque podem ser provenientes de atores estatais e não-estatais, utilizando uma multiplicidade de ferramentas, numa atividade de interferência, influência, operações ou guerra, dirigidas a um alvo. As ferramentas para retirar a capacidade de decisão do alvo podem ser ataques cinéticos contra infraestruturas, intrusão e exploração das infraestruturas em seu benefício, operações cobertas e clandestinas por parte dos serviços secretos hostis, investimentos estrangeiros em sectores estratégicos, atividade cibernética, campanha de desinformação através de órgãos de comunicação social controlados por empresas de fachada, exploração de lacunas ou fragilidades no aparelho judiciário e judicial, interferência e manipulação de grupos da sociedade para causar instabilidade e roturas na sociedade, entre tantas outras. Mas por entre tantas ferramentas, uma das mais poderosas é a desinformação, muito suportada pelas capacidades que a inteligência artificial lhe empresta.

De uma forma simples pode-se definir como desinformação a disseminação deliberada de informação falsa com o objetivo de manipular e influenciar a opinião pública. E esta ação tem como base três pilares. O primeiro é a manipulação emocional, onde a desinformação procura especialmente o sentido de raiva, medo ou indignação por forma a negar a utilização do espírito crítico por parte da audiência. O segundo é a amplificação algorítmica, onde os algoritmos utilizados pelas redes sociais preferem os conteúdos emocionais independentemente da sua veracidade. E por fim as fábricas de *bots*[3], onde redes organizadas de forma sistemática difundem massivamente desinformação através de milhares de contas

3 Os *bots* são programas de *software* que executam automaticamente tarefas específicas, frequentemente simulando o comportamento humano na interação com utilizadores, plataformas digitais ou outros sistemas informáticos.

falsas. Os poderes conferidos pela inteligência artificial podem servir inadvertidamente os interesses dos poderes hostis que representação a ameaça híbrida. Mas como é que se combate esta ameaça?

Uma possível solução passa por providenciar literacia digital e espírito crítico aos cidadãos, através do único possível antídoto existente atualmente: educação e treino desde tenras idades. A segunda poderá ser o desenvolvimento de instrumentos legais que permitam a deteção, proteção e erradicação de ameaças híbridas que intentem contra os direitos, liberdades e garantias dos cidadãos num Estado de direito. Quanto mais conhecimento existir sobre as táticas, técnicas e procedimentos dos atores estatais e não-estatais na utilização das "*armas*" híbridas para desestabilizar, interferir ou destruir os princípios democráticos, mais fácil se tornará o seu combate. Aqui a IA pode ter um papel importante na defesa contra estas ameaças híbridas, através de deteção de padrões suspeitos, confirmação da veracidade da informação, ou aconselhamento de ações defensivas para as inibir ou aniquilar.

IX. MICRO-DIRECIONAMENTO COGNITIVO

No âmbito do conceito de FIMI (*Foreign Information Manipulation and Interference*), definido pela União Europeia como atividades de manipulação da informação conduzidas de forma intencional e coordenada por atores estatais ou não estatais, incluindo os seus representantes ou intermediários (*proxies*), tanto no interior como no exterior dos respetivos territórios, o micro-direcionamento cognitivo constitui uma das ameaças mais significativas às democracias contemporâneas. Através da análise algorítmica dos perfis individuais, os cidadãos tornam-se alvos de campanhas de desinformação concebidas para moldar perceções, influenciar atitudes e alterar comportamentos.

Esta dinâmica pode gerar uma crescente dependência dos *bots* e de outros sistemas automatizados de interação, potenciando campanhas de manipulação de natureza política, moral, ideológica, religiosa ou até romântica. No plano político, uma das narrativas mais recorrentes procura convencer os indivíduos de que o Ocidente conduz uma guerra contra determinados valores, identidades ou culturas. Na perspetiva moral, difunde-se a ideia de que as democracias ocidentais são hipócritas, decadentes, reprimem os mais vulneráveis e apoiam regimes antidemocráticos sempre que isso serve os seus interesses estratégicos. No quadro

idealista, promove-se a perceção de que a democracia liberal é incapaz de proporcionar uma sociedade justa ou ideal, ao passo que modelos políticos assentes em valores russos, chineses ou islâmicos são apresentados como alternativas superiores. Numa perspetiva religiosa, sustenta-se que as democracias liberais e os respetivos ordenamentos jurídicos contradizem valores considerados tradicionais ou sagrados. Finalmente, existe um segmento da população particularmente recetivo a narrativas que romantizam a revolução, a oposição ao sistema e uma auto-proclamada superioridade moral, conferindo-lhes uma dimensão heroica.

Este novo paradigma informacional torna-se evidente através dos dados provenientes das plataformas digitais. O Facebook declarou ter identificado cerca de 5,8 mil milhões de contas falsas em 2022, sendo plausível admitir que o número real seja ainda superior. Esta realidade favorece um crescimento exponencial da desinformação, da informação incorreta e maliciosa. Algumas estimativas indicam que apenas 20 indivíduos conseguem criar a perceção de que foram publicadas entre 20.000 e 40.000 mensagens numa única hora. Nas eleições presidenciais norte-americanas de 2016, cerca de 120 contas do Facebook, responsáveis por aproximadamente 80.000 publicações, alcançaram um universo estimado de 126 milhões de pessoas. Acresce que se trata de uma estratégia caracterizada por uma relação custo-benefício particularmente elevada.

Paralelamente, os influenciadores digitais, mais do que os jornalistas, passaram a dominar plataformas como o TikTok e o Instagram enquanto fontes de informação para uma parte significativa da população mais jovem. Segundo diversas fontes, cerca de 29% dos jovens entre os 18 e os 24 anos obtêm notícias através do TikTok. Entre os utilizadores desta plataforma, 55% afirmam informar-se através de influenciadores ou outras personalidades digitais, enquanto apenas 33% recorrem a conteúdos produzidos por jornalistas. No Instagram, essas proporções situam-se nos 52% e 42%, respetivamente. Esta transformação dificulta cada vez mais a distinção entre informação factual, opinião e conteúdo manipulado.

Paralelamente, as notícias falsas propagam-se significativamente mais depressa do que as verdadeiras. Um estudo conduzido pelo *Massachusetts Institute of Technology* concluiu que as notícias falsas têm cerca de 70% mais probabilidade de serem partilhadas, sobretudo por despertarem reações emocionais mais intensas e maior curiosidade. O mesmo estudo demonstra que notícias verdadeiras raramente ultrapassam mil partilhas, enquanto conteúdos falsos podem atingir as 100 mil.

Com a evolução da inteligência artificial, o desafio já não consiste apenas em distinguir o que é verdadeiro do que é falso, mas também em determinar quem é efetivamente real. A crescente dificuldade em comprovar a autenticidade de pessoas, imagens, vídeos ou conteúdos digitais produz efeitos psicológicos relevantes, podendo gerar tanto medo e desconfiança como complacência e aceitação acrítica. A conjugação de realidades artificiais com técnicas de micro-direcionamento potenciadas pela inteligência artificial constitui, atualmente, uma das principais ameaças de manipulação e interferência da informação por agentes estrangeiros (FIMI) nos processos eleitorais.

Nas recentes eleições de um Estado-Membro da União Europeia, foi reportada uma campanha de desinformação conduzida por *bots* especificamente dirigida a jovens entre os 14 e os 16 anos. Não se pretendia influenciar diretamente o seu voto, uma vez que estes não possuíam idade legal para votar, mas antes induzi-los a influenciar os seus pais e avós, procurando alterar preferências eleitorais tradicionalmente mais conservadoras e associadas aos partidos políticos estabelecidos.

A subversão proporcionada pela inteligência artificial no domínio cognitivo poderá estar a transformar as tradicionais guerras de atrição em guerras cognitivas, conforme defende o general norte-americano David L. Goldfein. Neste contexto, o principal campo de batalha deixa de ser exclusivamente físico e passa a situar-se na esfera cognitiva, onde a consciência, a perceção e os processos de pensamento se tornam os principais alvos da ação estratégica. A influência deixa de incidir apenas sobre aquilo que as pessoas pensam, procurando moldar a própria forma como pensam. Consequentemente, a inteligência artificial tende a alterar profundamente a forma como os Estados exercem influência, coerção, subversão e conduzem a guerra. Simultaneamente, democratizou o acesso de atores não estatais a instrumentos de influência com alcance global.

Os rápidos avanços da inteligência artificial, das ciências cognitivas, das neuro-tecnologias e de outras áreas conexas aumentam igualmente o potencial para operações de manipulação em larga escala. Assiste-se, assim, à crescente militarização da mente, entendida como um novo espaço de conflito. Daí resulta a crescente utilização, no pensamento estratégico e militar, do conceito de domínio cognitivo, privilegiando formas de atuação não cinéticas em detrimento da força física convencional. Contudo, a ausência de um quadro internacional de governação suficientemente robusto cria oportunidades para que atores hostis explorem estas

capacidades, permanecendo significativas limitações nos mecanismos atualmente disponíveis para prevenir, detetar e contrariar estas formas de subversão no século XXI.

Neste contexto, o micro-direcionamento cognitivo potenciado pela inteligência artificial afirma-se como a nova fronteira do domínio cognitivo. O seu espectro estende-se desde a análise dos comportamentos digitais, tais como os cliques, pesquisas ou interações nas redes sociais, até à exploração das vulnerabilidades emocionais dos utilizadores. A questão central deixa de ser apenas quem influencia quem, passando igualmente por compreender de que forma os mecanismos psicológicos de validação, confirmação e dependência emocional podem ser explorados através de técnicas avançadas de segmentação comportamental. O risco converte-se numa ameaça concreta quando estas práticas produzem alterações no comportamento eleitoral ou comprometem a confiança dos cidadãos nos processos democráticos. Neste sentido, crescem as preocupações relativamente às futuras eleições europeias e aos processos eleitorais dos Estados-Membros da União Europeia. Entre os riscos mais relevantes encontra-se igualmente o enfraquecimento da legitimidade dos governos democraticamente eleitos, limitando a sua capacidade de atuação, particularmente em matérias de política externa e de defesa.

X. CONCLUSÕES

A Inteligência Artificial afirma-se como uma das forças mais transformadoras do século XXI, com impactos profundos nos domínios económico, social, educativo e militar. Contudo, esta transformação não é neutra nem isenta de riscos. Ao longo deste artigo, procurou-se demonstrar que, a par dos ganhos evidentes em eficiência, produtividade e capacidade de processamento, emergem ameaças menos visíveis, mas potencialmente mais disruptivas: a erosão da capacidade cognitiva individual, a fragilização das instituições e a crescente dependência de sistemas cuja lógica nem sempre é plenamente compreendida pelos seus utilizadores.

A automatização, quando não acompanhada por um espírito crítico robusto, pode conduzir a uma delegação excessiva do pensamento humano, transformando indivíduos em meros validadores de *outputs* algorítmicos. Este fenómeno, aliado à proliferação de desinformação amplificada por IA, cria condições propícias à manipulação social e à perda de confiança nas instituições, com implicações diretas na estabilidade democrática. No plano laboral, embora se perspetive um saldo líquido positivo na criação de emprego, a transição será exigente, impondo uma reconfiguração profunda das competências e dos sistemas educativos, que nem sempre acompanham o ritmo da mudança tecnológica.

No domínio da segurança e defesa, a integração da IA redefine o próprio conceito de poder militar, como evidenciado pela utilização massiva de drones e sistemas autónomos em conflitos recentes. Esta evolução levanta não apenas desafios operacionais, mas também dilemas éticos e jurídicos, sobretudo no que respeita à responsabilização e ao controlo humano sobre sistemas letais.

No que concerne à segurança nacional e à estabilidade democrática, a emergência do domínio cognitivo como campo de batalha revela que a desinformação potenciada por IA é uma das maiores ameaças híbridas da atualidade. A facilidade com que algoritmos e *bots* podem manipular as perceções individuais exige uma resposta que não seja apenas tecnológica, mas sobretudo pedagógica. Em última análise, o sucesso desta transição tecnológica não reside na substituição do humano pela máquina, mas na capacidade de os Estados e as sociedades promoverem a literacia digital e regulamentação como o antídoto principal contra a

manipulação, garantindo que o ser humano permaneça no centro do processo de decisão e da construção do conhecimento.

Perante este cenário, torna-se imperativo adotar uma abordagem equilibrada: potenciar os benefícios da IA sem abdicar da centralidade do ser humano. Isso implica investir seriamente na literacia digital, na educação crítica e em mecanismos de governação robustos. Mais do que substituir o homem, a IA deve ser concebida como um instrumento ao seu serviço e não como o seu substituto.

A resposta não deve passar por travar a inovação, que poderá trazer enormes benefícios, desde avanços médicos até soluções para desafios globais. Em vez disso, os governos devem garantir que os ganhos da IA sejam distribuídos de forma mais equilibrada. Medidas como uma tributação mais eficaz dos lucros extraordinários, bem como programas de requalificação profissional, podem ajudar os trabalhadores a adaptar-se. Se estas medidas forem insuficientes, poderão surgir propostas mais radicais, como a participação pública em empresas de IA ou a distribuição direta de parte dos seus lucros aos cidadãos. O colapso do emprego ainda não chegou, mas esperar por provas definitivas antes de reforçar a proteção social poderá revelar-se um erro. Os governos devem começar a preparar-se já.



O PROFISSIONAL DE SAÚDE NA ERA DIGITAL: RESPONSABILIDADE, AUTONOMIA E CARGA ADMINISTRATIVA

DANIELA MARQUES DIAS *

&

CARLA BARBOSA †

* Doutoranda em Direito Civil na Faculdade de Direito da Universidade de Coimbra. Assistente Convidada na Universidade de Leiria e Oeste. Bolseira de investigação do Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, no âmbito do Projeto *SafeCaring*. E-mail: daniela.dias@fd.uc.pt. ORCID: 0009-0001-0022-2346.

† Investigadora do Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra. Investigadora e associada do Centro de Direito Biomédico. E-mail: cbarbosa@fd.uc.pt. ORCID: 0000-0002-1845-9253.

RESUMO:

A Era Digital transformou, de forma significativa, o exercício do profissional de saúde, impondo uma reconfiguração de responsabilidades, deveres de diligência e a sua própria autonomia. Embora frequentemente associada à promessa de desburocratização e ao aumento de eficiência, esta metamorfose simboliza novas exigências jurídicas e éticas que impactam diretamente a prática clínica diária. Neste contexto, o profissional de saúde assume a posição de sujeito jurídico num ecossistema digital marcado pela interoperabilidade de sistemas, pela automação de processos e pela crescente utilização de ferramentas de apoio baseadas em inteligência artificial.

O presente artigo analisa criticamente estas dinâmicas, explorando a responsabilidade médica e administrativa em ambiente digital e o papel da supervisão humana na decisão assistida por IA. Sustenta que a automação não elimina a responsabilidade humana, reforçando, antes, a necessidade de verificação crítica e uma documentação rigorosa, primordiais para a salvaguarda do paciente e para a proteção jurídica do próprio profissional de saúde. São, igualmente, discutidos os riscos associados ao *automation bias*, evidenciando-se a relevância da autonomia crítica e da literacia digital na mitigação de decisões baseadas numa confiança excessiva nos sistemas. Ademais, expõe a carga administrativa como instrumento de conformidade legal e de segurança jurídica, destacando o risco de desresponsabilização na ausência de registos e o valor probatório da documentação clínica em contextos litigiosos. Paralelamente, aborda a tensão entre eficiência e ética no cuidado, por via da análise da intensificação burocrática como fator capaz de comprometer a relação terapêutica e de como processos de desburocratização não enquadrados podem significar uma fragilização na responsabilidade.

Por fim, propõe linhas de orientação para uma reforma jurídica sensível à prática clínica, assente na proporcionalidade documental, na proteção da autonomia profissional e na transparência dos sistemas digitais. Neste domínio, destaca-se o contributo do Projeto *SafeCaring* enquanto abordagem integradora para a construção de um modelo ético-jurídico sustentável, no qual a tecnologia apoia o cuidado, reforça a responsabilidade e preserva a primazia da relação humana nos cuidados de saúde.

Palavras-chave: Saúde Digital; Inteligência Artificial; Profissional de Saúde; Desburocratização; *SafeCaring*.

ABSTRACT:

The Digital Age has fundamentally restructured healthcare practice, imposing a reconfiguration of professional responsibilities, duties of care, and clinical autonomy. While digital transformation is often championed for its potential to reduce bureaucracy and enhance efficiency, it introduces complex legal and ethical mandates that reshape daily clinical operations. Within this digital ecosystem – characterized by interoperability, automation and AI-driven decision support – healthcare professionals emerge as distinct legal subjects. This article critically examines these dynamics, analysing medical and administrative liability in digital environments and the imperative of human oversight in AI-assisted care. It posits that automation, rather than absolving practitioners of responsibility, underscores the necessity of critical verification and rigorous documentation as pillars of patient safety and professional legal protection.

Furthermore, the study addresses the risks of automation bias, emphasizing the role of critical autonomy and digital literacy in mitigating over-reliance on automated systems. It investigates the administrative burden as a mechanism for legal compliance, highlighting the probative value of clinical records in litigation. Simultaneously, the text explores the ethical tension between efficiency and care, arguing that excessive bureaucratic can erode the therapeutic relationship, while unregulated debureaucratization may undermine accountability.

Finally, the article proposes guidelines for a legal reform grounded in documentary proportionality, professional autonomy, and systemic transparency. In this framework, the SafeCaring Project is presented as an integrative approach to ensure that technology supports care and reinforces responsibility while preserving the primacy of human relationships in healthcare.

Keywords: E-Healh; Artificial Intelligence; Health Professional; Debureaucratization; SafeCaring.

I. INTRODUÇÃO [1]

A digitalização dos sistemas de saúde[2] tem sido frequentemente apresentada como uma oportunidade[3] para simplificar procedimentos, reduzir burocracias e libertar o profissional para o cuidado.[4] Contudo, este movimento de desburocratização administrativa convive paradoxalmente com a emergência de novas exigências jurídicas, especialmente no domínio da responsabilidade, da proteção de dados, da supervisão humana sobre sistemas automatizados e, até mesmo, da auditabilidade das decisões clínicas. Daqui decorre que o profissional de saúde se encontra num espaço de tensão estrutural: por um lado, espera-se que execute menos tarefas administrativas, por outro, exige-se simultaneamente um nível mais elevado de registo, verificação e diligência técnica.

Esta ambivalência revela que a digitalização não implica necessariamente uma "desregulação", mas antes uma reconfiguração da regulação, em que a fronteira entre a simplificação e a intensificação burocrática se torna difusa. A redução de etapas formais pode coexistir com um acréscimo de obrigações documentais implícitas, necessárias para garantir segurança jurídica num ambiente marcado por interoperabilidade de sistemas, decisões assistidas por inteligência artificial[5] e fluxos automatizados. Como sujeito jurídico e ético, o

1 Este trabalho foi cofinanciado pelo Fundo Europeu de Desenvolvimento Regional (FEDER), através do Programa Centro2030, no âmbito do projeto nº 17358, "*Safecaring - Enhancing care safety and quality through reduced administrative burden*".

2 A COMISSÃO EUROPEIA define este fenómeno da saúde digital como "um conjunto de ferramentas e serviços que utilizam as tecnologias da informação e comunicação (TIC) para apoiar e melhorar todas as fases dos cuidados de saúde, desde a prevenção e o diagnóstico até ao tratamento, monitorização e gestão do estado de saúde", cf. https://health.ec.europa.eu/ehealth-digital-health-and-care/digital-health-and-care_pt (acedido a 18 de novembro de 2025).

3 Podendo, até, ser entendida como uma necessidade por alguns autores. Neste sentido, *vide* LUCA SEBANATI, et al., "Steps towards a digital health ecosystem", *Journal of Biomedical Informatics*, 44, 2011, p. 621, disponível para consulta em https://www.sciencedirect.com/science/article/pii/S1532046411000402 (acedido a 18 de novembro de 2025)

4 Estima-se que, no espaço de uma semana, médicos dediquem 19 horas a trabalho burocrático, o que contribui para o *burnout* destes profissionais, que terá consequências na qualidade dos cuidados de saúde prestados, cf. CHARLES MASON, "Bureaucracy in Healthcare", *PATIMES*, American Society for Public Administration, agosto de 2025, disponível para consulta em: https://patimes.org/bureaucracy-in-healthcare/ (acedido a 19 de novembro de 2025). *Vide*, ainda, MAFALDA FERREIRA DA SILVA, *A crise do Sistema Nacional de Saúde sob o olhar dos Médicos*, Dissertação de Mestrado de Gestão e Economia de Serviços de Saúde, Faculdade de Economia da Universidade do Porto, 2024, p. 21, disponível para consulta em: https://repositorio-aberto.up.pt/bitstream/10216/164557/2/702385.pdf (acedido a 19 de novembro de 2025); ADVISORY BOARD, *Physician burnout and depression, in 5 charts*, janeiro de 2024, disponível para consulta em: https://www.advisory.com/daily-briefing/2024/01/31/physician-burnout (acedido a 1 de dezembro de 2025).

5 Doravante IA.

profissional de saúde é chamado a atuar num quadro em que a promessa de eficiência convive com a obrigação de manter - e, por vezes, reforçar - os *standards* de diligência e autonomia clínica.

Neste sentido, apresentamos o *SafeCaring*[6], um projeto financiado pelo Fundo Europeu de Desenvolvimento Regional (FEDER), através do Programa Regional do Centro (CENTRO2030) do Portugal 2030, que envolve diversas instituições, nomeadamente a Universidade de Coimbra (Instituto Jurídico da Faculdade de Direito e Faculdade de Economia), o Instituto Pedro Nunes (IPN), *SafeCaring*, Pharmilab e o Hospital da Misericórdia da Mealhada. Trata-se de uma solução baseada em inteligência artificial que visa apoiar os profissionais de saúde na prestação de cuidados a utentes dependentes de forma mais personalizada, respeitando os tempos específicos de cada intervenção.

O *SafeCaring* pretende desenvolver um sistema integrado de apoio às intervenções clínicas, com o objetivo de reduzir significativamente o tempo despendido em registos pelos profissionais de saúde e minimizar efeitos adversos, como quedas, lesões de pressão e desidratação. Para tal, o sistema propõe alertar os profissionais sobre a necessidade de intervenções, utilizando técnicas de processamento de linguagem natural, e automatizar o registo das ações realizadas. A pedra basilar do projeto consiste, assim, na redução da carga administrativa, contribuindo para a melhoria da qualidade e da segurança dos cuidados prestados.

Neste contexto, o presente artigo procura analisar as implicações da digitalização na prática diária de cuidados de saúde, demonstrando como a desburocratização digital - quando não adequadamente enquadrada - pode gerar responsabilidades, criar vulnerabilidades jurídicas e afetar a autonomia profissional. A reflexão centra-se na forma como o profissional de saúde se reposiciona neste ecossistema digital, em que a simplificação administrativa não elimina - podendo até intensificar - a necessidade de uma ética sólida e de um quadro jurídico sensível às realidades clínicas, ao qual o SafeCaring poderá contribuir de forma significativa.

6 Para consultar o projeto, *vide* https://safecaring.pt/ .

II. O PROFISSIONAL DE SAÚDE COMO SUJEITO JURÍDICO NO ECOSSISTEMA DIGITAL

Um ecossistema natural consiste numa comunidade biológica de organismos ou agentes interativos, além do seu ambiente físico. Ora, aplicando esta definição ao domínio da saúde, podemos definir o ecossistema da saúde como a rede de uma infinidade de agentes: prestadores de cuidados (médicos, enfermeiros, farmacêuticos e outros profissionais de saúde), fornecedores de saúde, juntamente com as suas organizações e sistemas de informação, consumidores dos cuidados, além do ambiente socioeconómico e incluindo a estrutura institucional e regulatória da saúde.[7]

É possível transpormos o presente raciocínio para o digital, enquanto ecossistema transformador - e revolucionário - na forma como os profissionais de saúde exercem a sua atividade, geram e utilizam informação, interagem com pacientes e com os próprios sistemas. Assim, cremos que o profissional de saúde assume inevitavelmente, neste ecossistema digital, um papel de sujeito jurídico com deveres, direitos e responsabilidades específicos.[8] Esta condição jurídica implica considerar não apenas as regras tradicionais do exercício clínico, mas também as normas relacionadas com a proteção de dados pessoais e ferramentas que recorrem a inteligência artificial no seu funcionamento.

a. Responsabilidade médica e administrativa em ambiente digital

Apesar da precisão - por vezes superior à dos seres humanos - e do apoio que sistemas e produtos dotados de IA podem proporcionar na prestação de cuidados de saúde, a possibilidade de ocorrência de danos ao paciente não é eliminada. Destarte, revela-se imperativo abordar a responsabilidade médica no ambiente digital, que, à semelhança do que

7 Cf. LUCA SERBANATI, et al., "Steps towards...", *cit.*, p. 628; MEMED, "*O que é um ecossistema digital de saúde e qual sua razão de existir*", 2021, disponível para consulta em: https://blog.memed.com.br/ecossistema-digital-saude-memed/ (acedido a 2 de dezembro de 2025).

8 No mesmo sentido, *vide* LUCA SERBANATI, et al., "Steps towards...", *cit.*, p. 628; FEDERAÇÃO BRASILEIRA DE HOSPITAIS, "O papel dos médicos na evolução da medicina digital", *Revista Visão Hospitalar*, 2025, disponível para consulta em: https://revistavisaohospitalar.com.br/o-papel-dos-medicos-na-evolucao-da-medicina-digital/ (acedido a 2 de dezembro de 2025).

sucede na medicina tradicional, pode assentar na má prática médica, com base na violação das *leges artis* ou na violação do consentimento informado.[9]

É de reforçar que a inteligência artificial não substitui o juízo clínico, nem os profissionais de saúde, devendo funcionar apenas como um mero auxílio[10], um meio complementar, uma ferramenta de que aqueles dispõem no exercício das suas funções, respeitando as *leges artis*.[11] Ademais, o recurso a sistemas algorítmicos, por exemplo, não é o mesmo que delegar as suas funções nestes. Vejamos. O Código Deontológico da Ordem dos Médicos[12], no artigo 4.º, prevê os deveres de conduta do profissional, que terão de ser observados sob pena de violação das boas práticas médicas, e, nos termos do n.º 6 do artigo 11.º, "o médico não pode incumbir outros profissionais de saúde da realização de diagnóstico, prescrição ou gestão clínica autónoma de doentes"[13]. Entendemos que este último revela a intenção do legislador de ser exigível a atuação direta do médico na tomada de decisões e na execução dos procedimentos com o desígnio de salvaguardar os padrões éticos impostos pela profissão. Partindo deste exemplo, afigura-se-nos claro, portanto, que o profissional permanece responsável pela decisão final, devendo compreender as limitações desses algoritmos - nomeadamente, eventuais vieses ou resultados não plausíveis e que fornecem somente uma

9 Sobre o tema, *vide* GUILHERME DE OLIVEIRA, "*O fim da "arte silenciosa: o dever de informação dos médicos*", *Revista de Legislação e Jurisprudência*, Ano 128, n.º 3852, julho de 1995, pp. 103 e ss.

10 O estudo "Médicos Europeus e IA" realizado pelo Medscape concluiu que 75% dos médicos portugueses consideram que a utilização de IA na saúde - em especial, em radiologia e em clínica geral - é uma mais valia para o auxílio da sua atividade profissional e que irá contribuir para a redução de casos de negligência médica, cf. https://medjournal.pt/2024/11/14/9-em-cada-10-medicos-portugueses-defendem-um-enquadramento-legal-para-o-uso-da-inteligencia-artificial-no-ambiente-de-trabalho-medico/ (acedido a 3 de dezembro de 2025). No mesmo sentido, *vide* MICHELLE DIAS, *Inteligência Artificial na prestação dos cuidados de saúde e a perspetiva dos médicos portugueses*, XX Mestrado em Saúde Pública, Escola Nacional de Saúde Pública da Universidade Nova de Lisboa, 2020, pp. 42-44, disponível para consulta em: https://run.unl.pt/bitstream/10362/129610/1/RUN%20-%20Disserta%C3%A7%C3%A3o%20de%20Mestrado%20-%20Michelle%20Belleti%20Dias.pdf (acedido a 3 de dezembro de 2025).

11 Seja na escolha do tratamento como no diagnóstico médico, através, v.g. da análise e interpretação de exames imagiológicos. *Vide* entre outros, SHUROUG ALOWAIS, et al., "*Revolutioning healthcare: the role of artificial intelligence in clinical pratice*", *BMC Medical Education*, Volume 23, 2023, disponível para consulta em: https://link.springer.com/article/10.1186/s12909-023-04698-z#citeas (acedido a 18 de novembro de 2025)

12 Regulamento n.º 707/2016, de 21 de julho. Doravante CDOM.

13 Esta lógica é reforçada no Regulamento n.º 698/2019, de 5 de setembro (regulamento que define os atos profissionais próprios dos médicos, a sua responsabilidade, autonomia e limites). No artigo 5.º, n.º 3, é permitida a delegação de atos noutros profissionais de saúde, explicitando, porém, no n.º 2, que "*não deve incumbir outros profissionais de saúde da realização de atos próprios dos médicos*". É de ressalvar, porém, que, a nosso ver, o diploma faz coincidir a noção de ato com os deveres, o que se apreende com o texto legal do artigo 6.º, n.º 1: "o ato médico consiste na atividade diagnóstica, prognóstica, de vigilância (…), de prescrição e execução de medidas terapêuticas (…)."

correlação estatística - e readaptar a sua capacidade crítica dos resultados alcançados pelos algoritmos, uma vez que, nos termos do n.º 8 do artigo 4.º do CDOM, "*deve cuidar da permanente atualização da sua cultura científica e da sua preparação técnica.*"

A IA veio, ainda, simplificar tarefas administrativas[14], pelo que, o ambiente digital cria obrigações administrativas específicas, refletindo que o profissional adquire o estatuto de gestor de informação. Assim, assume-se, juntamente com as instituições, como responsável pelo tratamento adequado de dados de saúde, enquanto categoria especial de dados pessoais sujeita a uma proibição de tratamento pelo RGPD[15] e, em caso de violação, serão impostas sanções.

O profissional deve, assim, operar dentro das regras definidas pelas ordens profissionais e instituições.[16] A título exemplificativo, mencionamos a utilização obrigatória de sistemas de prescrição eletrónica, registos em plataformas autorizadas - evitando o armazenamento local indevido de dados e garantindo um acesso protegido (cf. artigos 36.º e 37.º do CDOM) - rastreabilidade e auditoria de atos digitais. A não observância destas normas pode resultar em responsabilidade administrativa, ainda que não tenha existido um dano físico clínico. Posto isto, a digitalização reforça a obrigação de documentar adequadamente todas as decisões, até porque, desta forma, se torna mais fácil a tarefa de averiguar quem realizou aquele ato.[17]

---

14 Sobre este ponto, *vide* FELIPE PENA/MAGDA FARIA, "Reflexões sobre o uso da Inteligência Artificial em ambientes de profissionais de saúde: impactos atuais e futuros", *Contemporary Journal*, Vol. 4, n.º 1, 2024, pp. 2840-2841, disponível para consulta em: https://ojs.revistacontemporanea.com/ojs/index.php/home/article/view/2908/2359 (acedido a 3 de dezembro de 2025); PETER PARYCEK, et al., "Artificial Intelligence (AI) and Automation in Administrative Procedures: Potentials, Limitations and Framework Conditions", *Journal of the Knowledge Economy*, Volume 15, 2024, pp. 8390-8415, disponível para consulta em: https://link.springer.com/article/10.1007/s13132-023-01433-3 (acedido a 5 de dezembro de 2024).

15 Regulamento n.º 679/2016, de 27 de abril relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados).

16 Cf. ANA LUÍS PEREIRA, "O papel da Ordem dos Médicos na transformação digital da saúde", *HealthNews*, 2025, disponível para consulta em: https://healthnews.pt/2025/05/14/o-papel-da-ordem-dos-medicos-na-transformacao-digital-da-saude/ (acedido a 2 de dezembro de 2025)

17 Neste sentido, o Regulamento (UE) 2025/327 do Parlamento Europeu e do Conselho, de 11 de fevereiro (relativo ao Espaço Europeu de Dados de Saúde) prevê que os sistemas de registos eletrónicos de saúde mantenham registos de acesso e uso de dados de saúde eletrónicos pessoais (cf. considerando 19).

b. Impacto das ferramentas digitais na autonomia profissional

A relação entre a tecnologia e a autonomia revela-se ambígua, na medida em que amplia a autonomia, mas também a pode reduzir. Quanto à ampliação da autonomia, como já fora mencionado, as ferramentas digitais podem facilitar o acesso rápido a informação clínica e científica, apoiando a prática baseada em evidência e num enorme volume de literatura e, por conseguinte, contribuir para uma maior precisão diagnóstica, otimização da triagem e priorização dos casos. No caso concreto do *SafeCaring*, esta manifesta-se na redução da carga administrativa e na autonomização de registos de cuidados de saúde, colocando à disposição dos profissionais de saúde mais tempo para a tomada de decisões clínicas informadas adaptadas às situações concretas.

Por outro lado, a autonomia profissional é limitada, desde logo, por um juízo de que estas são mais precisas - aspeto que será oportunamente desenvolvido. Parece-nos que esta limitação também ocorrerá quando as ferramentas emitam alertas frequentes e que podem interromper o normal desenrolar do trabalho do profissional - que acaba por os silenciar e colocar em risco a segurança do paciente[18] - ao automatizarem decisões ou as condicionarem - por exemplo, sistemas que bloqueiam prescrições fora de protocolo[19], ao incorporarem algoritmos com uma opacidade tal que as lógicas de decisão são totalmente incompreendidas pelo profissional. Por fim, não podemos ignorar a possibilidade de criarem dependência tecnológica, reduzindo a liberdade de abordagem clínica e crítica.

Além disso, a digitalização introduz uma nova forma de supervisão institucional, já que todos os acessos são registados, todas as decisões são auditáveis, os algoritmos podem emitir alertas e bloqueios automáticos e existe monitorização em tempo real da produtividade e conformidade. Isto transforma a autonomia profissional numa autonomia condicionada, em que o profissional mantém liberdade técnica, mas dentro de limites definidos por plataformas,

---

18 Cf. MARYATI YUSOF, "*The appropriateness of clinical decision support systems alerts in supporting clinical workflows: A systematic review*", *Health Informatics Journal*, Volume 27, n.º 2, 2021, p. 12, disponível em: https://www.researchgate.net/publication/350905000_The_appropriateness_of_clinical_decision_support_systems_alerts_in_supporting_clinical_workflows_A_systematic_review (acedido a 2 de dezembro de 2025)

19 Cf. SOPHIE BRONSTEIN, "*Using shared clinical decision support to reduce adverse drug events and improve patient safety*", *Frontiers in Digital Health*, dezembro de 2025, disponível em: https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2025.1703141/full (acedido a 2 de dezembro de 2025); LUCREZIA ARMANDO, et al., "*Clinical decision support systems to improve drug prescription and therapy optimisation in clinical practice: a scoping review*", *NMJ Health Care Inform*, Volume 30, n.º 1, maio de 2023, disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC10163516/

entidades reguladoras e sistemas de gestão. Para manter e reforçar a autonomia[20], apela-se à formação dos profissionais de saúde para que desenvolvam literacia digital[21] e, por conseguinte, competência de avaliação crítica de ferramentas tecnológicas, capacidade de interpretar os resultados fornecidos pelos algoritmos, conhecimento sobre dados e plataformas e, obviamente, competências éticas relacionadas com o uso de IA. Só desta forma se garantirá uma interação segura com a tecnologia, percecionando erros cometidos pelo sistema e evitando danos maiores para os pacientes.[22]

III. AUTOMAÇÃO E RESPONSABILIDADE: O PAPEL DA SUPERVISÃO HUMANA

Sem embargo das promessas de precisão diagnóstica, redução de erros humanos e otimização de recursos, a automação dos sistemas não elimina - e, em muitos casos, reforça - a necessidade de supervisão humana qualificada. O profissional de saúde mantém-se, como mencionado anteriormente, o sujeito jurídico central do processo assistencial, responsável por validar, interpretar e enquadrar as recomendações produzidas por algoritmos, sob pena de ser responsabilizado nos termos acima explicitados.[23] Assim, o papel da supervisão humana constitui um elemento essencial neste domínio com o objetivo de assegurar segurança e integridade ética no ecossistema digital da saúde.

---

20 Só através da salvaguarda da autonomia do profissional é possível evitar a substituição do mesmo e manter a ética na prestação de cuidados de saúde centrados no paciente, cf. MONIQUE HENDRIKS, et. Al, "Respecting Human Autonomy in Critical Care Clinical Decision Support", *Frontiers in Computer Science*, 2021, pp. 1-3, disponível para consulta em: https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2021.690576/full (acedido a 3 de dezembro de 2025).

21 Apesar de a amostra não ser circunscrita a médicos, afigura-se pertinente mencionar um estudo intitulado por "Consumer Sentiment Survey 2023" realizado pela *Boston Consulting Group*, que concluiu que 68% de portugueses manifestou a necessidade de mais formação prévia à utilização generalizada de IA em ambiente laboral, sendo que somente 13% da amostra recebeu, de facto, formação sobre a utilização de IA, cf. BUSSINESS.IT, disponível para consulta em: https://business-it.pt/2023/12/12/estudo-revela-que-23-dos-portugueses-utilizam-ferramentas-de-ia-%20semanalmente/ (acedido a 5 de dezembro de 2025).

22 No mesmo sentido, *vide* MARTA CANSADO, "Inteligência artificial no setor da saúde: desafios jurídicos e regulação", *Gabinete de Estratégia e Estudos*, 2024, p. 20, disponível para consulta em: https://www.gee.gov.pt/RePEc/WorkingPapers/GEE_PAPERS_187.pdf ; JOÃO FERREIRA, et al. "Empowering Health Professionals with Digital Skills to Improve Patient Care and Daily Workflows", *Healthcare*, Volume 13, n.º 3, 2025, disponível para consulta em: https://www.mdpi.com/2227-9032/13/3/329

23 Remetemos para o ponto II. a.

a. O enquadramento jurídico da decisão assistida por IA

Nos dias de hoje, a utilização de IA na saúde situa-se num quadro jurídico complexo, que combina normas de responsabilidade civil, legislação sobre a proteção de dados e, recentemente, regulamentos específicos sobre IA. Apesar da sofisticação tecnológica, o ordenamento jurídico europeu atual parte do princípio de que a IA assiste, mas *nunca substitui* o profissional[24], pelo que a decisão final é sempre imputada ao humano - o que se sucede, desde logo, por não se reconhecer personalidade jurídica a tais sistemas[25] e por, no caso dos profissionais de saúde, estes não a poderem delegar ao sistema, como explicado anteriormente.

Este princípio é reforçado pelo *AI Act*[26], que classifica a generalidade dos sistemas de IA aplicados à saúde como de alto risco[27]/[28], impondo, no artigo 14.º, uma obrigação de supervisão humana, ao mencionar que "*os sistemas de IA de risco elevado devem ser concebidos e desenvolvidos de modo a poderem (...) ser eficazmente supervisionados por pessoas singulares durante o período em que estão em utilização*" (n.º 1). Esta supervisão pretende "*prevenir ou minimizar os riscos para a saúde, a segurança ou os direitos*

24 Vide NÚCLEO DE SAÚDE DIGITAL GLOBAL E RELAÇÕES INTERNACIONAIS/UNIDADE DE CIBERSEGURANÇA, *White Paper, Inteligência Artificial na Saúde em Portugal: Regulamentação, Impactos e Perspetivas do Futuro*, fevereiro de 2025, p. 4, disponível para consulta em: https://www.spms.min-saude.pt/wp-content/uploads/2025/03/White-Paper_Inteligencia-Artificial-na-Saude-em-Portugal_-Final2-1.pdf (acedido a 2 de dezembro de 2025).

25 Sobre o tema, *vide* MAFALDA MIRANDA BARBOSA, *Inteligência Artificial - Entre a Utopia e a Distopia, Alguns Problemas Jurídicos*, GESTLEGAL, 1.ª edição, Coimbra, 2021, pp. 7-51.

26 Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, que cria regras harmonizadas em matéria de inteligência artificial e que altera os Regulamento (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 2018/858 e (UE) n.º 2018/1139 e (UE) n.º 2019/2144 e as Diretivas 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (Regulamento da Inteligência Artificial).

27 Explicaremos brevemente em que termos. O Regulamento da IA prevê duas condições cumulativas para que se classifique um sistema de alto risco: a previsão do produto na lista do Anexo I ao Regulamento e a sujeição do sistema à verificação de conformidade por parte de terceiros, de acordo com o artigo 6.º, n.º 1, a) e b). Dentro do primeiro critério, acresce que serão considerados como sistemas de alto risco os previstos no Anexo III, se consistirem num risco significativo de dano para a saúde, segurança ou direitos fundamentais das pessoas, segundo o artigo 6.º, n.º 2 e 3 do mencionado diploma. Assim, serão os produtores a ter de comprovar que, apesar de o seu sistema de incluir na previsão do Anexo III, este não representará um significativo risco para os fatores mencionados. Daqui decorre que, atendendo ao primeiro conjunto de critérios apresentados, os sistemas dotados de IA aplicados ao setor médico seriam, na sua maioria, tidos como sistemas de alto risco. O Anexo I consagra, em específico, os aparelhos médicos e os aparelhos para realização de diagnóstico in vitro, nos pontos 11 e 12. Assim, como seria expectável, desde logo, pela natureza sensível dos dados e da vulnerabilidade dos pacientes, estes sistemas estarão sujeitos a obrigações mais exigentes. Quanto ao segundo conjunto de critérios, é de atender ao Anexo III, que prevê os sistemas baseados em identificação biométrica e o sistema de triagem de pacientes nas urgências de hospitais (ponto 5, d).

28 Caso não se trate de um sistema de alto risco, é de mobilizar o argumento relativo à delegação de funções, na medida em que o profissional se manterá responsável pela interpretação que fizer do resultado fornecido pelo sistema - apesar de não o supervisionar - pois já o faria com os meios tradicionais.

fundamentais que possam surgir quando um sistema de IA de risco elevado é usado em conformidade com a sua finalidade prevista ou em condições de utilização indevida razoavelmente previsível, em especial quando esses riscos persistem apesar da aplicação de outros requisitos estabelecidos na presente secção", cf. n.º 2 do artigo 14.º.

Tratam-se, portanto, de sistemas que, pela sua natureza, não poderão funcionar sem supervisão humana adequada, impondo-se que sejam transparentes[29] [30] e, assim, explicáveis de forma compatível com o uso clínico, exigindo que o utilizador (*in casu*, o profissional de saúde) compreenda a lógica e as limitações da ferramenta - dentro dos possíveis, já que existem aspetos inteiramente incompreensíveis, até mesmo pelos próprios humanos que os programam (*blackbox effect*[31]).

Embora fabricantes e instituições que os utilizam (hospitais e clínicas) tenham responsabilidades (pelo *design*, colocação no mercado, segurança, manutenção/atualização), o profissional mantém responsabilidade clínica por decidir utilizá-las[32] e por, como mencionado, interpretar corretamente o resultado fornecido pelo sistema, respeitando as *leges artis* e garantindo que a decisão é centrada no paciente e que não foi delegada no

---

29 Cf. Artigo 13.º do AI Act. Esta exigência de transparência materializar-se-á em dois níveis: primeiro, quanto aos criadores dos sistemas de IA, procurando que sejam transparentes quanto ao uso de dados de saúde no treino dos sistemas; o segundo relativo aos pacientes, que deverão ter acesso a informação adequada quanto aos benefícios e desvantagens da utilização, cf. MINDY DUFFOURC/DOMINICK GIOVANNIELLO, "*The Autonomous AI Physician: Medical Ethics and Legal Liability*", *Multidisciplinary Perspectives on Artificial Intelligence and the Law,* Springer, 2024, pp. 207-228, disponível para consulta em: https://link.springer.com/chapter/10.1007/978-3-031-41264-6_11 (acedido a 4 de dezembro de 2025).

30 Em 2021, foi divulgado pela OMS o primeiro relatório global intitulado de *Ethics and Governance of Artificial Intelligence for Health* com o desígnio de estipular os princípios orientadores da IA: proteção da autonomia humana; garantia da transparência, explicabilidade e inteligibilidade; promoção de uma IA responsiva e sustentável; garantia da inclusão e equidade; fomento da responsabilidade e prestação de contas; promoção do bem-estar humano, segurança e interesse público. Disponível para consulta em: https://www.who.int/publications/i/item/9789240029200 (acedido a 2 de dezembro de 2025).

31 Tal como explicam MINDY DUFFOURC e DOMINICK GIOVANNIELLO, "*embora as funções algorítmicas utilizadas na análise de dados pela IA não sejam, e não possam ser, completamente transparentes por causa da "black box" da IA, os programadores devem, ainda assim, fornecer informações cruciais sobre a forma como a IA foi treinada e os potenciais enviesamentos do software para uma supervisão e análise independentes*", cf. "*The Autonomous AI Physician*..., *cit*., p. 213.

32 Numa publicação recente acerca os desafios de IA na saúde, a Associação Médica Americana estabeleceu diretrizes com três questões para que o médico avalie se deve utilizar/continuar a usar um sistema nas suas funções: primeiro, se a IA funciona (se se trata de uma tecnologia segura e efetiva com evidências científicas e capaz de respeitar os princípios éticos); se é viável para aquela prática clínica; e se o recurso a sistemas de IA traduzirá uma melhoria dos resultados clínicos e uma diminuição de danos para os pacientes. Cf. ELLIOTT CRIGER, et al, "*Trustworthy Augmented Intelligence in Health Care*", *Journal of Medical Systems*, 2022, pp. 12 e ss, disponível em: https://link.springer.com/article/10.1007/s10916-021-01790-z . *Vide*, ainda, RAFAELLA NOGAROLI, *Responsabilidade civil médica na Inteligência Artificial: culpa médica e deveres de conduta no século XXI*, Revista dos Tribunais, 2023, pp. 175-176.

algoritmo. Reforçamos que a decisão assistida continua ancorada no juízo humano, o que revela a responsabilidade do profissional de saúde caso adote cegamente uma recomendação automática.

b. Deveres de verificação, documentação e registo

O recurso a sistemas dotados de automação suscita o reforço dos deveres profissionais capazes de garantir a rastreabilidade, segurança e responsabilização, de entre os quais: o de verificação, de documentação e de registo. Relativamente ao *dever de verificação*, já num ambiente não digital, impõe-se que o profissional verifique se a recomendação algorítmica é coerente com os dados clínicos disponíveis, identifique eventuais inconsistências[33], vieses ou resultados improváveis, avalie se a ferramenta está atualizada e se foi usada em condições adequadas - o que, em última instância, se traduz na obrigação de supervisão que sobre si impende. A ausência de verificação - ou uma verificação meramente formal - pode configurar culpa - ou negligência - sobretudo quando o algoritmo apresenta limitações conhecidas ou documentadas[34].

Num ambiente não digital, o profissional de saúde tem o *dever de registar* os "*resultados que considere relevantes das observações clínicas*" (cf. n.º 1 do artigo 40.º CDOM) e o *dever de documentação* dos processos[35], decisões e dos critérios utilizados. Transpondo este dever para o domínio digital - e atendendo às suas especificidades - reforçamos a ideia de que cada interação com a ferramenta automatizada deve ser documentada explicitamente no processo eletrónico clínico, nomeadamente, a decisão feita com apoio do sistema, incluindo o caminho até a alcançar e que passa pela avaliação critica da recomendação, pelo motivo pelo qual foi seguida ou contrariada, eventuais riscos identificados e limitações tecnológicas que afetaram a decisão[36]. São estes dois últimos deveres que preservam a rastreabilidade, enquanto elemento central - e crucial - da digitalização da prestação de cuidados de saúde. Assim, cabe

33 Cf. MINDY DUFFOURC/DOMINICK GIOVANNIELLO, "*The Autonomous AI Physician*...", *cit*., p. 210.
34 De acordo com RAFAELLA NOGAROLI, trata-se de "*uma diligência mínima exigível a de que ele (o médico) afira de maneira crítica os resultados trazidos pelo sistema autónomo e, quanto mais grave a consequência de um erro nestas informações para o paciente, mais o médico deve questionar a base da decisão da Inteligência Artificial*.", cf. *Responsabilidade civil..., cit.*, p. 187.
35 Cf. ENTIDADE REGULADORA DA SAÚDE, *Deliberação do Conselho Diretivo da ERS*, 2007, p. 6, disponível em: https://www.ers.pt/uploads/writer_file/document/290/ERS-053-07.pdf (acedido a 2 de dezembro de 2025).
36 *Vide* CIRO MENNELLA, et al., "*Ethical and regulatory challenges of AI technologies in healthcare: a narrative review*", *Heliyon*, 10, 2024, disponível para consulta em: https://pmc.ncbi.nlm.nih.gov/articles/PMC10879008/ (acedido a 20 de novembro de 2025)

ao profissional garantir que os dados usados pelo algoritmo são devidamente tratados, como impõe o RGPD, através de anonimização e pseudonimização dos dados pessoais, que as entradas manuais não foram sujeitas a manipulação por pessoas não autorizadas e que o processo é passível de consulta futura, nomeadamente, em caso de litígio. Assim, a documentação e o registo corretos são primordiais tanto para a responsabilidade jurídica como para a melhoria contínua da utilização do sistema.

c. Automation bias e responsabilidade por confiança excessiva

O *automation bias* - definido como a tendência humana para confiar excessivamente nas decisões geradas por sistemas automatizados[37] - é um dos maiores riscos associados à IA em saúde. Este fenómeno manifesta-se de duas formas: erro por comissão e erro por omissão. O primeiro sucede quando quem toma a decisão adota medidas inadequadas com base em sugestões automatizadas que se revelam incorretas, sem as verificar em fontes de informação alternativas. Já o erro por omissão ocorre quando os tomadores de decisão não agem adequadamente por não terem sido informados ou alertados pelo sistema, mesmo quando havia sinais clínicos relevantes.[38]

Este fenómeno de *automation bias* aumenta nos casos em que o sistema é percebido como altamente preciso - mesmo não o sendo - quando o profissional tem excesso de carga de trabalho, nas hipóteses em que a decisão automatizada reduz a incerteza subjetiva e quando o algoritmo apresenta explicações convincentes - ainda que pouco transparentes. Todavia, a utilização de sistemas - e uma confiança excessiva nestes - não elimina a responsabilidade humana através do simples argumento de que a decisão foi do algoritmo, pelo que o profissional pode ser responsabilizado se seguir cegamente a recomendação sem a procurar compreender[39], se ignorar sinais clínicos contraditórios, se basear a sua decisão unicamente no sistema e quando não registar adequadamente o processo de decisão.

---

37 Cf. GIUSEPPE ROMEO/DANIELA CONTI, "*Exploring automation bias in human-AI collaboration: a review and implications for explainable AI*", *AI & Society*, 2025, pp. 4-5, disponível em: https://link.springer.com/article/10.1007/s00146-025-02422-7 (acedido a 6 de dezembro de 2025).

38 Cf. GIUSEPPE ROMEO/DANIELA CONTI, "Exploring automation...", *cit.*, p. 4.

39 As boas práticas impõem o dever mínimo de o médico procurar outro diagnóstico, seja em reunião com os pares ou reuniões multidisciplinares, com o objetivo de dissipar dúvidas, de acordo com o artigo 4.º da Convenção de Oviedo e o artigo 11º do CDOM, cf. SANDRA PASSINHAS, "*Comentário ao artigo 4.º da Convenção para a Proteção dos Direitos Humanos e da Dignidade do Ser Humano face às Aplicações da Biologia e da Medicina*", *Convenção para a Proteção dos Direitos do Homem e da Dignidade do Ser Humano*

Ademais, pode simbolizar o fenómeno *deskilling*[40], enquanto "*perda de perícia de competência dos profissionais que deixarão de exercitar a técnica e habilidade necessárias*"[41]. Assim, reforçamos que a formação dos profissionais de saúde adequada ao ecossistema digital contribua para se evitar o *automation bias* e responsabilização em caso de dano sofrido pelo paciente. Desta forma, a automação na saúde não elimina responsabilidade nem reduz exigências éticas, pelo que podemos afirmar que há uma deslocação do eixo da responsabilidade para a supervisão humana. É imperativo que o profissional seja capaz de compreender as limitações da IA, verificar e documentar criticamente cada recomendação, reconhecer e mitigar o *automation bias* e assegurar decisões seguras, explicáveis e juridicamente sustentáveis. Pelo que cremos que a automação só se torna verdadeiramente segura quando é acompanhada de autonomia crítica, transparência e responsabilidade humana ativa.

IV. CARGA ADMINISTRATIVA E CONFORMIDADE LEGAL

A digitalização dos sistemas de saúde não eliminou a burocracia, mas transformou-a. Os *dossiers* em papel estão a ser substituídos por plataformas eletrónicas, sistemas de registo e interfaces de prescrição que ampliaram a rastreabilidade e reforçaram as exigências de conformidade legal.

Esta carga administrativa, frequentemente percebida como onerosa pelos profissionais, cumpre funções essenciais no plano jurídico e organizacional. Contudo, a sua utilização inadequada, ausência ou preenchimento insuficiente pode gerar riscos sérios, desde a insegurança jurídica até à (des)responsabilização e fragilidade probatória em caso de litígio.

---

face a Aplicações da Biologia e da Medicina - 20 anos de Vigência em Portugal, Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2022, pp. 113-131.

40 *Vide* JOHN GROSSER, et al., "*Studying the Potential Effects of Artificial Intelligence on Physician Autonomy: Scoping Review*", *JMIR AI*, Volume 4, março de 2025, p. 6, disponível para consulta em: https://ai.jmir.org/2025/1/e59295/PDF (acedido a 5 de dezembro de 2025)

41 Cf. RUTE TEIXEIRA PEDRO, "*Breves reflexões sobre a reparação de danos causados na prestação de cuidados de saúde com utilização de robots*", *Inteligência Artificial e Robótica: Desafios para o Direito do Século XXI*, (coord. Sónia Moreira, Pedro Freitas), 1ª edição, Coimbra, GESTLEGAL, 2022, p. 158.

a. A burocracia como instrumento de segurança jurídica

Embora a burocracia seja, muitas vezes, vista como um obstáculo ao tempo clínico, constitui um instrumento de proteção jurídica[42], tanto para o paciente como para o profissional de saúde.

Os processos/registos clínicos e formulários eletrónicos - de consentimento, prescrições - servem para assegurar transparência no processo assistencial, garantir coerência entre atos clínicos e informação disponível, tornar rastreável cada decisão e permitir a auditoria e a supervisão da prática. A burocracia não existe apenas para controlo administrativo, mas também como um mecanismo de segurança jurídica, impedindo arbitrariedades e fornecendo um quadro claro de aferição de responsabilidades. Além disso, serve para proteção do profissional, uma vez que o registo completo e rigoroso o protege contrainterpretações - possivelmente erradas - de decisões clínicas, relata o cumprimento das *leges artis* e a diligência, bem como prova a adequação do comportamento em situações de risco ou incerteza[43]. Daqui decorre que, sem documentação, a boa prática pode tornar-se difícil ou impossível de demonstrar.

b. O risco de desresponsabilização na ausência de registos

A falta ou insuficiência de registos compromete a qualidade dos cuidados, aumenta o risco clínico - por quebrar a continuidade daqueles - e cria um ambiente de desresponsabilização difusa. Quando não existe documentação adequada, as decisões clínicas tornam-se invisíveis, passos importantes do processo de prestação de cuidados são perdidos, erros ou omissões não podem ser identificados e torna-se, assim, praticamente impossível reconstituir o percurso terapêutico do paciente. Esta ausência fragiliza o acompanhamento clínico e dificulta a continuidade de cuidados. Além disto, a ausência de registos cria, de certa forma, uma ausência de imputação clara de responsabilidade porque não é possível identificar quem decidiu, não se sabe quando a decisão foi tomada e não há

42 Na perspetiva clássica de MAX WEBER, a burocracia serve a segurança jurídica ao criar normas claras, previsíveis e rotinas padronizadas, de forma a alcançar estabilidade e racionalidade à administração pública e ao direito.

43 Neste sentido, *vide* MARCELLO FLORES, "*A importância dos registos médicos precisos para casos de responsabilidade médica*", *Jusbrasil*, outubro de 2023, disponível para consulta em: https://www.jusbrasil.com.br/artigos/a-importancia-dos-registros-medicos-precisos-para-casos-de-responsabilidade-medica/2006493624 (acedido a 2 de dezembro de 2025).

provas de que a informação foi corretamente analisada. Num contexto de equipas multidisciplinares e sistemas digitais interligados, a falta de documentação pode levar a conflitos de atribuição de culpa ou, pior, à impossibilidade de identificar o responsável, prejudicando o paciente e comprometendo a confiança no sistema.

A ausência de registos não é apenas um problema jurídico, mas um risco clínico direto e um dever dos profissionais de saúde[44]. Erros de medicação, falhas de diagnóstico e decisões descoordenadas ocorrem com maior frequência quando a informação não está documentada ou não é acessível.

c. Obrigações documentais e prova em caso de litígio

Num processo litigioso - civil, disciplinar, penal - os registos clínicos e administrativos constituem a principal prova.[45] Assim, os registos têm força probatória porque constituem documentação produzida no momento da ação, sendo considerados relatórios técnicos elaborados por profissionais qualificados que reforçam a autenticidade.[46] Em litígio, como mencionado, o que não está registado tende a ser considerado como não realizado.

As obrigações documentais - decorrentes dos deveres de documentação e de registo[47] - incluem o registo da história clínica, descrição de sintomas, decisões, anotações que podem revelar o raciocínio clínico e justificações, consentimento informado (incluindo modalidades digitais), utilização de ferramentas digitais e recomendações algorítmicas, comunicação com o paciente e a família, incidentes e eventos adversos.

Em contexto litigioso, a documentação serve para esclarecer a sequência dos factos, permitindo aos peritos médicos que auxiliam o juiz avaliar a conformidade com as boas práticas, determinar se ocorreu culpa, imprudência ou imperícia, protegendo o profissional quando demonstra diligência e zelo. Daqui decorre, logicamente, que um registo bem feito será determinante no afastamento da responsabilidade[48].

---

44 Previsto nos respetivos Códigos Deontológicos.

45 Cf. MANUELA MILLEN, "*A importância do prontuário médico na apuração da responsabilidade civil do médico*", Universidade Santa Cecília, 2020, pp. 51-52, disponível em: https://www.manuelamarcatti.com/wp-content/uploads/2023/10/Dissertacao_MANUELAMARCATTIVENTURADECAMARGOMILLEN471.pdf

46 Cf. BRANDY PATRICK, "*When a Medical Record Becomes a Legal Document*", *ABA*, 2025, disponível em: https://www.americanbar.org/groups/law_practice/resources/law-technology-today/2025/when-a-medical-record-becomes-a-legal-document/

47 Remetemos para o ponto III. b.

48 Neste sentido, *vide* NUNO PINTO OLIVEIRA, *Ilicitude e Culpa na Responsabilidade Médica*, Edição (I)Materiais para o Direito da Saúde, n.º 1, Centro de Direito Biomédico, Coimbra, 2019, pp. 86-87, disponível

Em suma, podemos afirmar que, apesar das críticas relativas à carga administrativa, esta desempenha funções estruturais no ecossistema jurídico e clínico digital, uma vez que assegura a continuidade assistencial, fornece segurança jurídica ao permitir a defesa de decisões em contexto litigioso e, assim, contribui para a proteção de ambas as partes na relação clínica. Assim, entendemos que a digitalização deve, idealmente, tornar estes processos mais eficientes, sem, contudo, os eliminar, pois constituem o alicerce normativo e probatório em saúde.

V. ENTRE A EFICIÊNCIA E A ÉTICA NO CUIDADO

Por tudo o que já fora mencionado, são percetíveis os novos desafios da digitalização na gestão do tempo clínico e na relação terapêutica. Ora, por um lado, espera-se que o profissional seja eficiente, produtivo e rigoroso no cumprimento de obrigações legais e documentais, enquanto, por outro, se mantém a exigência ética de um cuidado centrado no paciente, atento à dimensão humana, relacional e comunicacional da prática clínica. É nesta complexa tarefa de equilíbrio que se encontra, atualmente, um dos maiores dilemas do exercício profissional no mundo digital.

a. O equilíbrio entre tempo para o paciente e exigências legais

Pelo caminho percorrido, é inegável que a prática clínica digitalizada é marcada por uma tensão permanente entre duas necessidades: tempo para cuidar e tempo para registar e documentar, ambas essenciais, mas, por vezes, de difícil harmonização. Isto pode gerar redução do tempo disponível para ouvir e atender ao que o paciente transmite, espelhando uma quebra na continuidade da narratividade, alteração na dinâmica da consulta, centrando-a em ecrãs[49] e sistemas e erguendo uma sensação de despersonalização na interação - algo que deve ser evitado.

---

para consulta em: https://centrodedireitobiomedico.org/data/publicacoes/PUB2024826194943.pdf (acedido a 2 de dezembro de 2025).

49 Cf. ONUR ASAN, et al., "*More screen time, less face time – implications for EHR design*", *Journal of Evaluation in Clinical Practice*, Volume 20, issue 6, 2014, pp. 896-901, disponível em: https://onlinelibrary.wiley.com/doi/10.1111/jep.12182

Se, por um lado, as exigências legais visam proteger o paciente e o profissional de saúde, garantem segurança jurídica e asseguram padrões de qualidade, por outro, o seu cumprimento impera que seja retirada uma percentagem do tempo dedicado ao processo assistencial. No entanto, o tempo clínico não é apenas duração, mas sim qualidade da relação. Do ponto de vista ético, ouvir, observar e dialogar são atos fundamentais da prestação de cuidados de saúde. Dispor de tempo suficiente possibilita uma melhor compreensão dos sintomas e necessidades concretas. A confiança terapêutica constrói-se em interações capazes de gerar um espaço de informação e reflexão[50]. Quando o tempo para o paciente é reduzido, a autonomia do doente e a qualidade diagnóstica podem ser comprometidas, bem como um aumento do risco de desumanização da relação.

Posto isto, a conformidade legal é indispensável, assim como o tempo para o paciente. A nosso ver, o equilíbrio entre estes dois aspetos será exequível através da simplificação de processos redundantes, da integração de sistemas que evitem a duplicação de registos e do desenvolvimento de ferramentas de IA - como no projeto *SafeCaring* - que apoiem o profissional sem invadir a esfera relacional.

b. O impacto da desburocratização na qualidade da relação terapêutica

A nosso ver, a desburocratização[51] - quando bem implementada - é uma oportunidade para fortalecer a relação terapêutica, apresentando-se, ainda, como uma medida de justiça

50 *Vide* LINNI GU, et al., "*Patient perception of doctor communication skills and patient trust in rural primary health care: the mediating role of health service quality*", *BMC Primary Care*, Volume 23, n.º 255, disponível para consulta em: https://link.springer.com/article/10.1186/s12875-022-01826-4 (acedido a 5 de dezembro de 2025); AMÉRICO VARELA, *Comunicação e Decisão Partilhada com o Doente*, Dissertação de Mestrado em Gestão e Avaliação de Tecnologias em Saúde, Escola Superior de Tecnologia da Saúde de Lisboa, Instituto Politécnico de Lisboa, 2016, disponível em: https://repositorio.ipl.pt/entities/publication/44d63005-872b-410f-993e-65ff1ac2f82d (acedido a 5 de dezembro de 2025); SERKAN DENIZ, et al., "*The mediating role of shared decision-making in the Effect of the patient-physician relationship on compliance with treatment*", *Journal of Patient Experience*, Volume 8, 2021, disponível em: https://journals.sagepub.com/doi/epub/10.1177/23743735211018066 (acedido a 5 de dezembro de 2025).

51 De acordo com a INSPEÇÃO-GERAL DAS ATIVIDADES EM SAÚDE, "*desburocratizar implica retirar a carga burocrática, e no contexto da melhoria dos serviços, reduzir a complexidade administrativa e a rigidez da cadeia de coordenação e controlo de uma organização com o objetivo de simplificar e tornar mais rápidos e eficientes os procedimentos utilizados*", cf. IGAS, "Desburocratização e simplificação administrativa no Serviço Nacional de Saúde", *Memorando do Plano de Auditoria*, 2021, p. 19, disponível para consulta em: https://www.igas.min-saude.pt/wp-content/uploads/2023/11/AUD_Memo_Desburocratizacao_e_Simplificacao_Administrativa_no_SNS.pdf (acedido a 28 de novembro de 2025). *Vide*, ainda, BERT ROCKMAN. "*Bureaucracy*", *Encyclopedia Britannica*, julho de 2021, disponível para consulta em: https://www.britannica.com/topic/bureaucracy (acedido a 28 de novembro de 2025)

social.[52] Devendo, todavia, ser ponderada, já que reduzir ou eliminar etapas de verificação pode acarretar riscos clínicos e jurídicos que, de outra forma, não se verificariam.

No que diz respeito aos potenciais benefícios da desburocratização[53], é de apontar que a redução de tarefas administrativas confere ao profissional tempo para comunicar com o paciente, ouvindo-o, o que, por sua vez, fortalece o vínculo terapêutico, melhora a satisfação de ambas as partes da relação, reduz o *burnout* associado ao peso burocrático e favorece decisões clínicas mais ponderadas e com menos interrupções[54]. Ora, se é dado ao profissional mais tempo para concentrar a atenção no paciente - e na atividade assistencial - a consulta torna-se mais segura e eficaz e supera-se a possível despersonalização da relação capaz de abalar a confiança e o humanismo que caracterizam a mesma.

Por outro lado, é de apontar os riscos associados a uma *desburocratização excessiva*, tais como a eliminação de certos registos ou procedimentos, que pode gerar falta de rastreabilidade (com os problemas mencionados supra), uma fragilidade probatória em caso de litígio e a perda de informação essencial para a continuidade de cuidados, que pode criar um risco de erro clínico. Assim, a desburocratização não pode significar a ausência de documentação, devendo traduzir-se numa mais frutífera e menos redundante[55], que poderá ser concretizada através do recurso a tecnologia.

Cremos que uma desburocratização equilibrada favorece a dimensão ética do cuidado, dado que devolve ao profissional a capacidade de estar presente, promove uma comunicação mais empática e humanizada, reforça a confiança do paciente no sistema e permite àquele o exercício de uma autonomia e de um juízo crítico, superando a tradicional submersão em

---

52 Cf. MÁRIO PEIXOTO, "*Burocracia na Saúde*", *Diário do Minho*, fevereiro de 2025, disponível para consulta em: https://www.diariodominho.pt/opiniao/2025-02-08-burocracia-na-saude-67a6436c64486 (acedido a 28 de novembro de 2025).

53 Em Portugal, as tentativas de desburocratização iniciaram-se com a criação da Rede Interministerial de Modernização Administrativa (RIMA) - Decreto-Lei n.º 4/97, de 9 de janeiro. Mais recentemente, a Portaria n.º 11/2024, de 18 de janeiro veio reforçar a simplificação administrativa.

54 Neste sentido, *vide* ASSOCIAÇÃO PORTUGUESA DE MEDICINA GERAL E FAMILIAR, *Um Novo Futuro para a Medicina Geral e Familiar,* 2023, pp. 26 e ss, disponível para consulta em: https://apmgf.pt/wp-content/uploads/2023/05/Um-Novo-Futuro.pdf (acedido a 21 de novembro de 2025); CARLA PEREIRA, *O Tempo para Gerir. Como?*, Dissertação de Mestrado em Direção e Chefia de Serviços em Enfermagem, Escola Superior de Enfermagem do Porto, 2019, pp. 79-84, disponível para consulta em: https://comum.rcaap.pt/entities/publication/f58b0e27-52f2-46b6-9ff0-be38125da8c0 (acedido a 23 de novembro de 2025).

Atendendo aos dados de um estudo realizado na Universidade do Porto, concluiu-se que os próprios profissionais - neste caso, médicos - veem a desburocratização como uma prioridade, cfr. MAFALDA FERREIRA DA SILVA, *A crise..., cit.*, p. 34, 37, 44.

55 Cf. ASSOCIAÇÃO PORTUGUESA DE MEDICINA GERAL E FAMILIAR, *Um Novo Futuro..., cit.*, p. 26.

tarefas mecânicas. Tal não significa o completo abandono da burocracia, mas uma reconfiguração desta enquanto instrumento de apoio e não de sobrecarga.

Posto isto, afirmamos que o contacto entre a eficiência e a ética no cuidado se centra na forma como o sistema organiza o tempo e as obrigações do profissional. A digitalização pode aumentar a eficiência ou deteriorar a relação terapêutica – tudo depende da forma como os processos são concebidos. O desígnio final não é eliminar a burocracia, mas garantir que a documentação auxilia o cuidado prestado, que a tecnologia liberta tempo em vez de o capturar, que o profissional permanece disponível para o paciente e, claro está, que o vínculo humano continue no centro da prática clínica. É, precisamente, neste equilíbrio, que reside a verdadeira ética da saúde digital.

VI. CONCLUSÃO

É inquestionável que a transformação digital no setor de saúde trouxe novas possibilidades de eficiência, segurança e qualidade, tendo, ainda, introduzido desafios significativos para a autonomia profissional, a responsabilidade jurídica e a relação clínica. No ecossistema digital, o profissional de saúde é simultaneamente agente clínico, gestor de informação - tornando-se sujeito jurídico responsável pelos dados e decisões - utilizador crítico de tecnologias e ator regulado por sistemas digitais. Pelo que se encontra no cruzamento entre exigências legais crescentes, sistemas algorítmicos cada vez mais complexos e a necessidade de manter um cuidado centrado na pessoa.

A análise desenvolvida ao longo deste artigo revela que a digitalização jamais poderá dispensar da supervisão humana - antes a reforça - e que a burocracia, apesar de ser, de um ponto de vista, uma enorme carga administrativa, constitui um elemento estruturante da segurança clínica e jurídica, apelando-se, todavia, à sua simplificação, evitando-se uma desburocratização excessiva, e a uma formação que confira literacia digital a quem recorre às tecnologias, em especial, às dotadas de IA.

Neste sentido, torna-se indispensável uma reforma jurídica sensível à prática clínica, que reconheça a realidade concreta do trabalho dos profissionais e evite soluções tecnocráticas que ampliem encargos sem atender a uma melhoria do cuidado prestado. Assim, cremos que uma reforma eficaz se deve orientar por uma proporcionalidade documental, distinguindo-se

o que é essencial do que é redundante, privilegiando registos clínicos relevantes e eliminando requisitos meramente formais. Ademais, deve ser garantida pela legislação a proteção da autonomia profissional, assegurando que os sistemas digitais complementam - e não condicionam - o juízo clínico e evitando que o profissional se torne num mero operador de interfaces. Por fim, deveria ser valorizado pelas normas o tempo clínico de qualidade, incentivando tecnologias que o libertem para a relação terapêutica, ao invés de a substituírem.
É precisamente neste horizonte que se inscreve o contributo do *SafeCaring*, enquanto abordagem conceptual e metodológica orientada para a criação de ecossistemas digitais seguros, éticos e harmonizados. O projeto *SafeCaring* propõe uma articulação equilibrada entre tecnologia, responsabilidade e cuidado, oferecendo importantes contributos. Começando pela integração ético-jurídica desde o desenho tecnológico, evitando que sistemas digitais sejam implementados sem previamente serem avaliados os impactos sobre a autonomia profissional, a segurança, a privacidade e a carga administrativa. Envolve, ainda, modelos de supervisão humana qualificada, que reconhecem o papel insubstituível do profissional na interpretação de recomendações algorítmicas e na mitigação de riscos como o *automation bias*. Por fim, oferece um foco no valor relacional do cuidado, promovendo ferramentas que libertem tempo, reduzam complexidade e reforcem a confiança entre pacientes, profissionais e instituições.
A saúde digital necessita, portanto, não apenas de inovação tecnológica, mas também de uma inovação normativa que respeite os ritmos e exigências da prática clínica. Uma reforma jurídica alinhada com estes princípios - e inspirada nos contributos de *SafeCaring* - permitirá construir um sistema mais seguro, transparente e humano, onde a tecnologia serve o cuidado.

REFERÊNCIAS BIBLIOGRÁFICAS

ADVISORY BOARD, *Physician burnout and depression, in 5 charts*, janeiro de 2024. Disponível em: https://www.advisory.com/daily-briefing/2024/01/31/physician-burnout (acedido a 1 de dezembro de 2025)

ALOWAIS, Shuroug, et al., "*Revolutioning healthcare: the role of artificial intelligence in clinical practice*", *BMC Medical Education*, Volume 23, 2023. Disponível em: https://link.springer.com/article/10.1186/s12909-023-04698-z#citeas (acedido a 18 de novembro de 2025)

ARMANDO, Lucrezia, et al., "*Clinical decision support systems to improve drug prescription and therapy optimisation in clinical practice: a scoping review*", *NMJ Health Care Inform*, Volume 30, n.º 1, maio de 2023. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC10163516/ (acedido a 2 de dezembro de 2025)

ASAN, Onur, et al., "*More screen time, less face time - implications for EHR design*", *Journal of Evaluation in Clinical Practice*, Volume 20, issue 6, 2014, pp. 896-901. Disponível em: https://onlinelibrary.wiley.com/doi/10.1111/jep.12182 (acedido a 1 de dezembro de 2025)

ASSOCIAÇÃO PORTUGUESA DE MEDICINA GERAL E FAMILIAR, *Um Novo Futuro para a Medicina Geral e Familiar,* 2023. Disponível em: https://apmgf.pt/wp-content/uploads/2023/05/Um-Novo-Futuro.pdf (acedido a 21 de novembro de 2025)

BARBOSA, Mafalda Miranda, *Inteligência Artificial - Entre a Utopia e a Distopia, Alguns Problemas Jurídicos*, GESTLEGAL, 1.ª edição, Coimbra, 2021

BRONSTEIN, Sophie, "*Using shared clinical decision support to reduce adverse drug events and improve patient safety*", *Frontiers in Digital Health*, dezembro de 2025. Disponível em: https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2025.1703141/full (acedido a 2 de dezembro de 2025)

BUSINESS.IT, 2023. Disponível em: https://business-it.pt/2023/12/12/estudo-revela-que-23-dos-portugueses-utilizam-ferramentas-de-ia-%20semanalmente/ (acedido a 5 de dezembro de 2025)

CANSADO, Marta, "*Inteligência artificial no setor da saúde: desafios jurídicos e regulação*", *Gabinete de Estratégia e Estudos*, 2024. Disponível em: https://www.gee.gov.pt/RePEc/WorkingPapers/GEE_PAPERS_187.pdf

CRIGER, Elliott, et al, "*Trustworthy Augmented Intelligence in Health Care*", *Journal of Medical Systems*, 2022. Disponível em: https://link.springer.com/article/10.1007/s10916-021-01790-z (acedido a 4 de dezembro de 2025)

DENIZ, Serkan, et al., "*The mediating role of shared decision-making in the Effect of the patient-physician relationship on compliance with treatment*", Journal of Patient Experience, Volume 8, 2021. Disponível em: https://journals.sagepub.com/doi/epub/10.1177/23743735211018066 (acedido a 5 de dezembro de 2025)

DIAS, Michelle, *Inteligência Artificial na prestação dos cuidados de saúde e a perspetiva dos médicos portugueses*, XX Mestrado em Saúde Pública, Escola Nacional de Saúde Pública da Universidade Nova de Lisboa, 2020. Disponível em: https://run.unl.pt/bitstream/10362/129610/1/RUN%20-%20Disserta%C3%A7%C3%A3o%20de%20Mestrado%20-%20Michelle%20Belleti%20Dias.pdf (acedido a 3 de dezembro de 2025)

DUFFOURC, Mindy / GIOVANNIELLO, Dominick, "*The Autonomous AI Physician: Medical Ethics and Legal Liability*", *Multidisciplinary Perspectives on Artificial Intelligence and the Law,* Springer, 2024, pp. 207-228. Disponível em: https://link.springer.com/chapter/10.1007/978-3-031-41264-6_11 (acedido a 4 de dezembro de 2025)

ENTIDADE REGULADORA DA SAÚDE, *Deliberação do Conselho Diretivo da ERS*, 2007. Disponível em: https://www.ers.pt/uploads/writer_file/document/290/ERS-053-07.pdf (acedido a 2 de dezembro de 2025)

FEDERAÇÃO BRASILEIRA DE HOSPITAIS, "*O papel dos médicos na evolução da medicina digital*", *Revista Visão Hospitalar*, 2025. Disponível em: https://revistavisaohospitalar.com.br/o-papel-dos-medicos-na-evolucao-da-medicina-digital/ (acedido a 2 de dezembro de 2025)

FERREIRA, João, et al. "*Empowering Health Professionals with Digital Skills to Improve Patient Care and Daily Workflows*", *Healthcare*, Volume 13, n.º 3, 2025. Disponível em: https://www.mdpi.com/2227-9032/13/3/329 (acedido a 3 de dezembro de 2025)

FLORES, Marcello, "*A importância dos registos médicos precisos para casos de responsabilidade médica*", *Jusbrasil*, outubro de 2023. Disponível em: https://www.jusbrasil.com.br/artigos/a-importancia-dos-registros-medicos-precisos-para-casos-de-responsabilidade-medica/2006493624 (acedido a 2 de dezembro de 2025)

GROSSER, John, et al., "*Studying the Potential Effects of Artificial Intelligence on Physician Autonomy: Scoping Review*", *JMIR AI*, Volume 4, março de 2025. Disponível em: https://ai.jmir.org/2025/1/e59295/PDF (acedido a 5 de dezembro de 2025)

GU, Linni, et al., "*Patient perception of doctor communication skills and patient trust in rural primary health care: the mediating role of health service quality*", *BMC Primary Care*, Volume 23, n.º 255. Disponível em: https://link.springer.com/article/10.1186/s12875-022-01826-4 (acedido a 5 de dezembro de 2025)

HENDRIKS, Monique, et. al, "*Respecting Human Autonomy in Critical Care Clinical Decision Support*", *Frontiers in Computer Science*, 2021. Disponível em: https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2021.690576/full (acedido a 3 de dezembro de 2025)

INSPEÇÃO-GERAL DAS ATIVIDADES EM SAÚDE, "*Desburocratização e simplificação administrativa no Serviço Nacional de Saúde*", *Memorando do Plano de Auditoria*, 2021. Disponível em: https://www.igas.min-saude.pt/wp-content/uploads/2023/11/AUD_Memo_Desburocratizacao_e_Simplificacao_Administrativa_no_SNS.pdf (acedido a 28 de novembro de 2025)

MASON, Charles, "*Bureaucracy in Healthcare*", *PATIMES*, American Society for Public Administration, agosto de 2025. Disponível em: https://patimes.org/bureaucracy-in-healthcare/ (acedido a 19 de novembro de 2025)

MEDJOURNAL, 2024. Disponível em: https://medjournal.pt/2024/11/14/9-em-cada-10-medicos-portugueses-defendem-um-enquadramento-legal-para-o-uso-da-inteligencia-artificial-no-ambiente-de-trabalho-medico/ (acedido a 3 de dezembro de 2025)

MEMED, "*O que é um ecossistema digital de saúde e qual sua razão de existir*", 2021. Disponível em: https://blog.memed.com.br/ecossistema-digital-saude-memed/ (acedido a 2 de dezembro de 2025)

MENNELLA, Ciro, et al., "*Ethical and regulatory challenges of AI technologies in healthcare: a narrative review*", *Heliyon*, 10, 2024. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC10879008/ (acedido a 20 de novembro de 2025)

MILLEN, Manuela, *A importância do prontuário médico na apuração da responsabilidade civil do médico*, Universidade Santa Cecília, 2020. Disponível em: https://www.manuelamarcatti.com/wp-content/uploads/2023/10/Dissertacao_MANUELAMARCATTIVENTURADECAMARGOMILLEN471.pdf (acedido a 1 de dezembro de 2025)

NOGAROLI, Rafaella, *Responsabilidade civil médica na Inteligência Artificial: culpa médica e deveres de conduta no século XXI*, Revista dos Tribunais, 2023

NÚCLEO DE SAÚDE DIGITAL GLOBAL E RELAÇÕES INTERNACIONAIS / UNIDADE DE CIBERSEGURANÇA, *White Paper, Inteligência Artificial na Saúde em Portugal: Regulamentação, Impactos e Perspetivas do Futuro*, fevereiro de 2025. Disponível em: https://www.spms.min-saude.pt/wp-content/uploads/2025/03/White-Paper_Inteligencia-Artificial-na-Saude-em-Portugal_-Final2-1.pdf (acedido a 2 de dezembro de 2025)

OLIVEIRA, Guilherme de, "*O fim da "arte silenciosa: o dever de informação dos médicos*", *Revista de Legislação e Jurisprudência*, Ano 128, n.º 3852, julho de 1995 - pp. 70-72, 101-104

OLIVEIRA, Nuno Pinto, *Ilicitude e Culpa na Responsabilidade Médica*, Edição (I)Materiais para o Direito da Saúde, n.º 1, Centro de Direito Biomédico, Coimbra, 2019. Disponível em: https://centrodedireitobiomedico.org/data/publicacoes/PUB2024826194943.pdf

OMS, *Ethics and Governance of Artificial Intelligence for Health – WHO Guidance,* 2021. Disponível em: https://www.who.int/publications/i/item/9789240029200 (acedido a 2 de dezembro de 2025)

PARYCEK, Peter, et al., "*Artificial Intelligence (AI) and Automation in Administrative Procedures: Potentials, Limitations and Framework Conditions*", *Journal of the Knowledge Economy*, Volume 15, 2024, pp. 8390-8415. Disponível em: https://link.springer.com/article/10.1007/s13132-023-01433-3 (acedido a 5 de dezembro de 2024)

PASSINHAS, Sandra, "*Comentário ao artigo 4.º da Convenção para a Proteção dos Direitos Humanos e da Dignidade do Ser Humano face às Aplicações da Biologia e da Medicina*", *Convenção para a Proteção dos Direitos do Homem e da Dignidade do Ser Humano face a Aplicações da Biologia e da Medicina - 20 anos de Vigência em Portugal*, Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2022, pp. 113-132

PATRICK, Brandy, "*When a Medical Record Becomes a Legal Document*", *ABA*, 2025. Disponível em: https://www.americanbar.org/groups/law_practice/resources/law-technology-today/2025/when-a-medical-record-becomes-a-legal-document/ (acedido a 1 de dezembro de 2025)

PEDRO, Rute Teixeira, "*Breves reflexões sobre a reparação de danos causados na prestação de cuidados de saúde com utilização de robots*", *Inteligência Artificial e Robótica: Desafios para o Direito do Século XXI*, (coord. Sónia Moreira, Pedro Freitas), 1ª edição, Coimbra, GESTLEGAL, 2022, pp. 151-185

PEIXOTO, Mário, "*Burocracia na Saúde*", *Diário do Minho*, fevereiro de 2025. Disponível em: https://www.diariodominho.pt/opiniao/2025-02-08-burocracia-na-saude-67a6436c64486 (acedido a 28 de novembro de 2025)

PENA, Felipe/FARIA, Magda, "*Reflexões sobre o uso da Inteligência Artificial em ambientes de profissionais de saúde: impactos atuais e futuros*", *Contemporary Journal*, Vol. 4, n.º 1, 2024, pp. 2833-2846. Disponível em: https://ojs.revistacontemporanea.com/ojs/index.php/home/article/view/2908/2359 (acedido a 3 de dezembro de 2025)

PEREIRA, Ana Luís, "*O papel da Ordem dos Médicos na transformação digital da saúde*", *HealthNews*, 2025. Disponível em: https://healthnews.pt/2025/05/14/o-papel-da-ordem-dos-medicos-na-transformacao-digital-da-saude/ (acedido a 2 de dezembro de 2025)

PEREIRA, Carla, *O Tempo para Gerir. Como?*, Dissertação de Mestrado em Direção e Chefia de Serviços em Enfermagem, Escola Superior de Enfermagem do Porto, 2019, pp. 79-84. Disponível em: https://comum.rcaap.pt/entities/publication/f58b0e27-52f2-46b6-9ff0-be38125da8c0 (acedido a 23 de novembro de 2025)

ROCKMAN, Bert, "Bureaucracy", *Encyclopedia Britannica*, julho de 2021. Disponível em: https://www.britannica.com/topic/bureaucracy

ROMEO, Giuseppe/CONTI, Daniela, "*Exploring automation bias in human-AI collaboration: a review and implications for explainable AI*", *AI & Society*, 2025. Disponível em: https://link.springer.com/article/10.1007/s00146-025-02422-7 (acedido a 6 de dezembro de 2025)

SEBANATI, Luca, et al., "*Steps towards a digital health ecosystem*", *Journal of Biomedical Informatics*, 44, 2011. Disponível em https://www.sciencedirect.com/science/article/pii/S1532046411000402 (acedido a 18 de novembro de 2025)

SILVA, Mafalda Ferreira da, *A crise do Sistema Nacional de Saúde sob o olhar dos Médicos*, Dissertação de Mestrado de Gestão e Economia de Serviços de Saúde, Faculdade de Economia da Universidade do Porto, 2024. Disponível em: https://repositorio-aberto.up.pt/bitstream/10216/164557/2/702385.pdf (acedido a 19 de novembro de 2025)

VARELA, Américo, *Comunicação e Decisão Partilhada com o Doente*, Dissertação de Mestrado em Gestão e Avaliação de Tecnologias em Saúde, Escola Superior de Tecnologia da Saúde de Lisboa, Instituto Politécnico de Lisboa, 2016. Disponível em: https://repositorio.ipl.pt/entities/publication/44d63005-872b-410f-993e-65ff1ac2f82d (acedido a 5 de dezembro de 2025)

YUSOF, Maryati, "*The appropriateness of clinical decision support systems alerts in supporting clinical workflows: A systematic review*", *Health Informatics Journal*, Volume 27, n.º 2, 2021. Disponível em: https://www.researchgate.net/publication/350905000_The_appropriateness_of_clinical_decision_support_systems_alerts_in_supporting_clinical_workflows_A_systematic_review (acedido a 2 de dezembro de 2025)

LEGISLAÇÃO

Regulamento (UE) 2025/327 do Parlamento Europeu e do Conselho, de 11 de fevereiro relativo ao Espaço Europeu de Dados de Saúde

Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, que cria regras harmonizadas em matéria de inteligência artificial (Regulamento da Inteligência Artificial)

Regulamento n.º 698/2019, de 5 de setembro (Regulamento que define os atos profissionais próprios dos médicos, a sua responsabilidade, autonomia e limites)

Regulamento n.º 707/2016, de 21 de julho (Regulamento de Deontologia Médica)

Regulamento n.º 679/2016, de 27 de abril relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados)

Decreto-Lei n.º 4/97, de 9 de janeiro (Cria a Rede Interministerial de Modernização Administrativa (RIMA)

CYBERLAW
BY CIJIC

O ENSINO DO DIREITO NA ERA DA INTELIGÊNCIA ARTIFICIAL: A ÉTICA E A DECISÃO HUMANA COMO FUNDAMENTOS DO USO PEDAGÓGICO DA IA, A PARTIR DA EXPERIÊNCIA DA HISTÓRIA DO DIREITO

RAIMUNDO CHAVES NETO*

* Cidadão luso-brasileiro, é Professor Auxiliar Convidado da Faculdade de Direito da Universidade de Lisboa (FDUL) e investigador do Centro de Investigação Interdisciplinar IURIS (FDUL). Doutor em História do Direito pela FDUL (2025), Mestre em Direito Romano pela Universidade de Coimbra (2011) e especialista em Direitos Humanos pelo *Ius Gentium Conimbrigae* (2009), é profissional da educação desde 1993.
As atividades pedagógicas de integração da Inteligência Artificial descritas neste estudo são aplicadas pelo autor não apenas nas disciplinas histórico-jurídicas da FDUL, mas também nas disciplinas dogmáticas que rege na Universidade Europeia, designadamente Direitos Reais e Direito da Família. O presente estudo constitui um resultado intermédio de um Projeto de Investigação de Pós-Doutoramento em curso no Instituto de Ciências Histórico-Jurídicas da FDUL, sob a orientação do Professor Doutor Eduardo Vera-Cruz Pinto e em articulação com o Centro de Investigação Jurídica do Ciberespaço (CIJIC), dedicado ao uso pedagógico da Inteligência Artificial no ensino do Direito. O modelo pedagógico aqui exposto concorre ao Prémio de Inovação Pedagógica de 2026 da Universidade de Lisboa.
Todas as fontes legais, doutrinárias e institucionais foram verificadas pelo autor nas respetivas fontes oficiais, em coerência com a própria prática de verificação que o estudo propõe.
Contacto: raimundoneto@fd.ulisboa.pt.
O autor agradece ao Diretor da revista *Cyberlaw by CIJIC*, Dr. Nuno Teixeira Castro, o amável convite para colaborar com este texto.

RESUMO:

A difusão massiva dos modelos de linguagem de grande escala, a partir de novembro de 2022, colocou, nas mãos de todos os estudantes de Direito, instrumentos capazes de redigir textos jurídicos, sintetizar doutrina, traduzir fontes latinas e medievais e simular argumentação. Este estudo defende que o ensino do Direito será mais bem desenvolvido com a utilização criteriosa da Inteligência Artificial (IA), desde que essa utilização permaneça subordinada a dois fatores decisivos: a ética e a decisão humana. A IA não substitui o professor; potencia-o. Não decide; auxilia a decidir. Não interpreta as fontes; multiplica o acesso a elas e liberta tempo para a interpretação, que é tarefa irredutivelmente humana. O argumento constrói-se a partir das disciplinas histórico-jurídicas, o Direito Romano e a História do Direito Português, tomadas como laboratório privilegiado: disciplinas de fontes, de contexto e de juízo, cuja propensão a expor a alucinação e o anacronismo dos modelos as torna, não as mais ameaçadas pela IA, mas as mais bem armadas para a domesticar. Fixam-se cinco princípios éticos operativos e demonstra-se, com quatro guiões de aula já experimentados, que o fator decisivo do sucesso pedagógico não é a ferramenta, mas o desenho humano do seu uso. Conclui-se projetando o modelo, no quadro de um Projeto de Pós-Doutoramento, para a totalidade do currículo jurídico.

Palavras-chave: Inteligência Artificial; ensino do Direito; História do Direito; Direito Romano; ética; decisão humana; integridade académica; humanidades digitais.

ABSTRACT:

The mass diffusion of large language models since November 2022 has placed in the hands of every law student tools capable of drafting legal texts, synthesising doctrine, translating Latin and medieval sources and simulating legal argument. This paper argues that legal education will be better served by the judicious use of Artificial Intelligence (AI), provided that such use remains subordinate to two decisive factors: ethics and human decision. AI does not replace the teacher; it empowers the teacher. It does not decide; it assists decision. It does not interpret the sources; it multiplies access to them and frees time for interpretation, which remains an irreducibly human task. The argument is built from the historical-legal disciplines, Roman Law and the History of Portuguese Law, taken as a privileged laboratory: disciplines of sources, context and judgement, whose tendency to expose the hallucination and anachronism of the models makes them not the most threatened by AI but the best equipped to tame it. Five operative ethical principles are set out, and it is shown, through four classroom scripts already tested, that the decisive factor of pedagogical success is not the tool but the human design of its use. The paper concludes by projecting the model, within a Post-Doctoral Project, onto the whole of the legal curriculum.

Keywords: Artificial Intelligence; legal education; History of Law; Roman Law; ethics; human decision; academic integrity; digital humanities.

I. INTRODUÇÃO: O PROBLEMA E A TESE

O ensino universitário do Direito atravessa uma transformação tecnológica sem precedente desde a invenção da imprensa. A difusão massiva, a partir de novembro de 2022, dos modelos de linguagem de grande escala (*large language models*, doravante LLM) colocou nas mãos de todos os estudantes instrumentos capazes de redigir textos jurídicos, sintetizar doutrina, traduzir fontes latinas e medievais, simular argumentação e responder a perguntas de exame.[1] Esta realidade não é uma hipótese de futuro: é o quotidiano das salas de aula, que o autor conhece diretamente como docente de Direito Romano e de História do Direito Português da Faculdade de Direito da Universidade de Lisboa (FDUL).

Perante este facto, três atitudes são possíveis. A primeira é a proibição, que a experiência demonstra ineficaz e pedagogicamente empobrecedora, além de incoerente com um mundo profissional (o foro, a advocacia, os tribunais, a administração pública) que adota aceleradamente estas ferramentas. A segunda é a adoção acrítica, que degrada o ensino, dissolve a autoria, fragiliza a avaliação e ameaça precisamente aquilo que a formação jurídica visa constituir: o juízo crítico do jurista, a capacidade de interpretar fontes no seu contexto, a consciência de que o Direito não se reduz à lei nem a *outputs* de sistemas computacionais. A terceira atitude, que este estudo defende, é a integração crítica, eticamente fundada e pedagogicamente orientada, da IA no ensino do Direito, mantendo a decisão humana (do professor e do estudante) como centro irredutível do processo de ensino e aprendizagem.

A tese a demonstrar é, por isso, simultaneamente afirmativa e limitativa: afirmativa, porque sustenta que o ensino do Direito será mais bem desenvolvido com a utilização da IA; limitativa, porque condiciona esse melhoramento à primazia da ética e da decisão humana. Não há aqui paradoxo, mas a própria lição da História do Direito: tal como o *ius* romano não se confundia com a *lex*, e a *iuris prudentia* não se deixou reduzir a um catálogo de soluções, também o ensino jurídico não pode confundir-se com a transmissão automatizada de informação. A máquina dá respostas; o professor ensina a fazer perguntas.

---

1 Sobre a arquitetura que sustenta estes modelos, veja-se Vaswani, A., et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30; sobre o desempenho em provas jurídicas e, em simultâneo, as suas limitações, Choi, J. H., Hickman, K. E., Monahan, A. B., & Schwarcz, D. (2022). ChatGPT goes to law school. *Journal of Legal Education*, 71(3), pp. 387-400, e Katz, D. M., Bommarito, M. J., Gao, S., & Arredondo, P. (2024). GPT-4 passes the bar exam. *Philosophical Transactions of the Royal Society A*, 382(2270), 20230254.

Este estudo constrói o argumento a partir de um campo determinado, as disciplinas histórico-jurídicas, e projeta-o, na parte final, para a totalidade do currículo jurídico. A escolha metodológica justifica-se: o que nas disciplinas de fontes se vê com nitidez, porque nelas o erro do modelo é imediatamente verificável contra um texto de autoridade, vale, com as devidas adaptações, para o Direito Civil, o Direito Penal, o Direito Administrativo ou o Direito da União Europeia, onde a fonte a confrontar é a lei vigente, o acórdão ou a doutrina dominante. O plano da exposição é o seguinte: justifica-se a eleição da História do Direito como laboratório (II); fixa-se o enquadramento normativo e ético do uso da IA no ensino superior (III); desenvolvem-se os dois fundamentos da tese, a ética (IV) e a decisão humana (V); inventariam-se as ferramentas por função pedagógica (VI); demonstra-se a operatividade do modelo com quatro guiões de aula já experimentados (VII); trata-se o ponto nevrálgico da avaliação e da integridade académica (VIII) e as exigências de segurança, verificação de fontes e propriedade intelectual (IX); e projeta-se, enfim, o modelo para todas as disciplinas jurídicas, no quadro de um Projeto de Pós-Doutoramento (X), antes das conclusões (XI). Fica fora do objeto, salvo na medida do enquadramento, a utilização da IA na decisão judicial, matéria de literatura própria, e a regulação geral da IA enquanto tal.

II. PORQUÊ A HISTÓRIA DO DIREITO COMO LABORATÓRIO

A eleição da História do Direito como campo de aplicação não é casual nem meramente biográfica. As disciplinas histórico-jurídicas, o Direito Romano e a História do Direito Português na tradição curricular da FDUL, são, por três razões convergentes, o laboratório ideal para pensar o uso pedagógico da IA no Direito.

Em primeiro lugar, são disciplinas de fontes. O seu ensino assenta no contacto com textos: a Lei das XII Tábuas, o *Corpus Iuris Civilis*, as Ordenações do Reino, a Lei da Boa Razão, os forais, as cartas de povoação. Ora, é precisamente no tratamento de grandes massas textuais, pesquisa, transcrição, tradução, indexação e comparação, que a IA e as humanidades digitais oferecem ganhos imediatos e mensuráveis, como demonstram instrumentos consolidados há décadas para as fontes romanísticas e para a transcrição automática de manuscritos.[2] A

2 Para as fontes romanísticas, veja-se o software *Amanuensis*, de Riedlberger, P., & Rosenbaum, G., construído sobre a base ROMTEXT da Universidade de Linz; para a transcrição automática de manuscritos, Kahle, P.,

tradição da chamada Escola de Lisboa, que faz da edição crítica e do contacto direto com a fonte a base da pedagogia histórico-jurídica, encontra na digitalização das fontes um prolongamento natural, e não uma rutura.[3]

Em segundo lugar, são disciplinas de contexto e de juízo. A História do Direito ensina que toda a norma tem um tempo, um espaço, uma sociedade e uma intencionalidade; que interpretar é situar; que o anacronismo é o erro metodológico capital. António Manuel Hespanha fez desta advertência contra a leitura teleológica do passado jurídico um eixo da renovação da historiografia jurídica portuguesa.[4] Os modelos de linguagem, treinados sobre dados maioritariamente contemporâneos e propensos a alucinações e a anacronismos, falham exatamente onde a História do Direito é mais exigente. Esta tensão é pedagogicamente fecunda: usar a IA nas disciplinas histórico-jurídicas obriga o estudante a verificar, a confrontar com as fontes, a detetar o erro, isto é, a fazer aquilo que o jurista sempre fez perante qualquer texto de autoridade duvidosa. A IA, bem usada, converte-se num exercício permanente de crítica de fontes.

Em terceiro lugar, a História do Direito fornece o quadro conceptual para pensar a própria IA jurídica. A distinção romana entre *auctoritas* e *potestas*, entre *ius* e *lex*, entre a criação jurisprudencial do Direito e a sua imposição legislativa, eixos do programa de Direito Romano, ilumina de modo singular o debate atual sobre decisão algorítmica, juiz-robô e codificação do Direito.[5] Quem estudou a degradação da *iuris prudentia* perante a concentração normativa do *princeps*, e a substituição do Direito dos juristas pelo Direito dos códigos, dispõe de instrumentos críticos únicos para avaliar a promessa, e o perigo, de um Direito calculado por máquinas. Neste sentido, o estudo inverte a relação habitual: não é apenas a IA que serve a História do Direito; é a História do Direito que serve a compreensão crítica da IA.

Colutto, S., Hackl, G., & Mühlberger, G. (2017). Transkribus: A service platform for transcription. *ICDAR*, pp. 19-24.

3 Fouto, A. C., Rodrigues Rocha, F., Seixas, M., Brigas, M. A., & Silva Santos, J. (2018). *Textos de história do direito português*. AAFDL. A coletânea, organizada por docentes do grupo de Ciências Histórico-Jurídicas da FDUL, é adotada como base de trabalho nas aulas práticas e constitui o acervo controlado sobre o qual se desenham, na secção VII, os ambientes de estudo ancorados.

4 Hespanha, A. M. (2012). *Cultura jurídica europeia: Síntese de um milénio*. Almedina. A advertência metodológica contra as leituras teleológicas do passado jurídico vale, por inteiro, contra as narrativas deterministas do futuro tecnológico.

5 Vera-Cruz Pinto, E. (2022). *Introdução ao direito romano: As questões fundamentais*. AAFDL; e Vera-Cruz Pinto, E. (2020). *Curso de direito romano* (2.ª ed.). Principia.

Esta eleição exige, porém, que se situe o estudo no estado da arte, e que dele se destaque a lacuna que ocupa. O encontro entre a Inteligência Artificial e o Direito não é recente. A pergunta fundadora é de Alan Turing, em 1950, sobre se as máquinas podem pensar; a expressão *artificial intelligence* nasce na proposta do seminário de Dartmouth, de 1955.[6] No Direito, o marco inaugural é o artigo de Lee Loevinger sobre a jurimetria, de 1949, seguido da especulação pioneira de Bruce Buchanan e Thomas Headrick sobre raciocínio jurídico computacional e, nas décadas de 1970 e 1980, dos sistemas periciais jurídicos que Richard Susskind teorizou.[7] A viragem estatística do século XXI, da IA simbólica para a aprendizagem automática sobre grandes *corpora*, produziu a previsão computacional de decisões judiciais e a síntese de referência de Kevin Ashley, ao mesmo tempo que a reflexão crítica, de Mireille Hildebrandt a Frank Pasquale, alertava para o fim do Direito como o conhecemos e para a sociedade da caixa negra.[8] A terceira fase, a atual, é a da IA generativa: a arquitetura *transformer* e o lançamento público dos assistentes conversacionais universalizaram capacidades inéditas, documentadas em estudos sobre o desempenho dos modelos em provas jurídicas, ao mesmo tempo que a investigação media, em escala preocupante, as suas alucinações.[9] Este duplo registo, potência e falibilidade, é o ponto de partida realista de qualquer reflexão pedagógica séria.

Paralelamente, o campo da Inteligência Artificial na educação tem mais de meio século, dos sistemas tutores inteligentes aos atuais tutores generativos, e uma reflexão madura que vai do relatório *Intelligence Unleashed* às obras de síntese de Wayne Holmes e colegas, passando pela interrogação crítica de Neil Selwyn sobre se os robôs deveriam substituir os professores,

6 Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), pp. 433-460.

7 Loevinger, L. (1949). Jurimetrics: The next step forward. *Minnesota Law Review*, 33(5), pp. 455-493; e, sobre a teorização dos sistemas periciais jurídicos, Susskind, R., cuja reflexão se prolonga até *Tomorrow's lawyers* (3.ª ed., 2023).

8 Ashley, K. D. (2017). *Artificial intelligence and legal analytics*. Cambridge University Press; Aletras, N., Tsarapatsanis, D., Preoţiuc-Pietro, D., & Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights. *PeerJ Computer Science*, 2, e93; Hildebrandt, M. (2015). *Smart technologies and the end(s) of law*. Edward Elgar; e Pasquale, F. (2015). *The black box society*. Harvard University Press, bem como Pasquale, F. (2020). *New laws of robotics*. Belknap Press.

9 Vaswani, A., et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30; Choi, J. H., Hickman, K. E., Monahan, A. B., & Schwarcz, D. (2022). ChatGPT goes to law school. *Journal of Legal Education*, 71(3), pp. 387-400; Katz, D. M., Bommarito, M. J., Gao, S., & Arredondo, P. (2024). GPT-4 passes the bar exam. *Philosophical Transactions of the Royal Society A*, 382(2270), 20230254; e, sobre as alucinações, Dahl, M., Magesh, V., Suzgun, M., & Ho, D. E. (2024). Large legal fictions. *Journal of Legal Analysis*, 16(1), pp. 64-93.

a que respondeu que não.[10] A irrupção da IA generativa deslocou o debate para a integridade académica, a avaliação e a literacia, e fez da orientação centrada no ser humano o seu critério partilhado, como ilustram a orientação da UNESCO e a proposta de Ethan Mollick do humano ao leme.[11] A investigação empírica começa a distinguir a IA que amplifica o esforço cognitivo da IA que o substitui, a chamada descarga cognitiva, distinção operativa também para este estudo.

A reflexão sobre tecnologia e ensino jurídico, que tem em Susskind o seu autor mais influente, concentrou-se, porém, nas disciplinas dogmáticas e na formação profissionalizante, deixando por cartografar as disciplinas de formação cultural e crítica do jurista.[12] As humanidades digitais aplicadas à história jurídica, por seu turno, produziram instrumentos notáveis, mas sem reflexão pedagógica sistemática; e a literatura sobre IA na educação ainda não pensou especificamente as disciplinas de formação histórica e cultural dos currículos profissionalizantes. O cruzamento específico, a IA e o ensino da História do Direito, está por fazer, em Portugal e, em larga medida, internacionalmente. É essa a lacuna que o estudo ocupa, e é a partir dela que se projeta, na parte final, para a totalidade do ensino jurídico.

III. ENQUADRAMENTO NORMATIVO E ÉTICO DO USO DA IA NO ENSINO SUPERIOR

O uso da IA no ensino superior português não é um espaço de indiferença jurídica. Convoca, desde logo, a Constituição da República Portuguesa: a liberdade de aprender e ensinar (artigo 43.º), a autonomia universitária (artigo 76.º, n.º 2), o direito à educação e à cultura (artigos 73.º e seguintes) e a garantia do artigo 35.º sobre utilização da informática, que proíbe decisões exclusivamente automatizadas lesivas de direitos sem garantias

10 Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed*. Pearson; Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education*. Center for Curriculum Redesign; Holmes, W., & Porayska-Pomsta, K. (Eds.) (2023). *The ethics of artificial intelligence in education*. Routledge; e Selwyn, N. (2019). *Should robots replace teachers?* Polity.

11 Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO; e Mollick, E. (2024). *Co-intelligence: Living and working with AI*. Portfolio/Penguin.

12 Susskind, R. (2023). *Tomorrow's lawyers* (3.ª ed.). Oxford University Press. Em contraponto, sobre a educação em geral, Selwyn, N. (2019). *Should robots replace teachers?* Polity.

adequadas, princípio que o Regulamento Geral sobre a Proteção de Dados densifica no seu artigo 22.º.[13]

No plano europeu, o Regulamento (UE) 2024/1689 (Regulamento Inteligência Artificial) constitui o primeiro quadro normativo geral do mundo sobre IA.[14] Para o ensino relevam, em especial, quatro disposições. O artigo 4.º impõe um nível suficiente de literacia em IA ao pessoal dos responsáveis pela implantação, o que abrange as universidades enquanto utilizadoras, e transforma a própria compreensão crítica da ferramenta em dever jurídico. O Anexo III, ponto 3, qualifica como de risco elevado os sistemas de IA usados para determinar o acesso ao ensino, avaliar resultados da aprendizagem ou detetar comportamentos proibidos de estudantes durante provas. O artigo 14.º exige supervisão humana efetiva dos sistemas de risco elevado. E o artigo 5.º proíbe práticas como o reconhecimento de emoções em instituições de ensino. Daqui resulta, como se verá, a exclusão deliberada, do modelo que se propõe, dos sistemas de *proctoring* algorítmico e de deteção automática de fraude.

Ao lado do Regulamento, mantêm relevância o Regulamento (UE) 2016/679 (RGPD), aplicável ao tratamento de dados pessoais de estudantes, e, no plano do Conselho da Europa, a Convenção-Quadro sobre Inteligência Artificial e Direitos Humanos, Democracia e Estado de Direito (2024) e a Carta Ética Europeia sobre a utilização da IA nos sistemas judiciais (CEPEJ, 2018), cujos cinco princípios, respeito pelos direitos fundamentais, não discriminação, qualidade e segurança, transparência e controlo pelo utilizador, são transponíveis, *mutatis mutandis*, para o ensino do Direito.[15]

No plano interno, além da Lei n.º 27/2021, de 17 de maio (Carta Portuguesa de Direitos Humanos na Era Digital), cujo artigo 9.º consagra o direito ao desenvolvimento de competências digitais, relevam a Estratégia Nacional de Inteligência Artificial e os instrumentos de *soft law* académica: as recomendações da UNESCO (a Recomendação sobre a Ética da Inteligência Artificial, de 2021, e a orientação para a IA generativa na educação e

13 Regulamento (UE) 2016/679 (RGPD), artigo 22.º, sobre decisões individuais automatizadas; e Constituição da República Portuguesa, artigos 35.º, 43.º, 73.º e seguintes, e 76.º, n.º 2.

14 Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024 (Regulamento Inteligência Artificial). *Jornal Oficial da União Europeia*, L, 2024/1689. Relevam em especial os artigos 4.º (literacia em IA), 5.º (práticas proibidas) e 14.º (supervisão humana), e o Anexo III, ponto 3 (sistemas de risco elevado em educação).

15 Comissão Europeia para a Eficiência da Justiça (CEPEJ) (2018). *European ethical charter on the use of artificial intelligence in judicial systems and their environment*. Conselho da Europa; e Conselho da Europa (2024). *Convenção-Quadro sobre Inteligência Artificial e Direitos Humanos, Democracia e Estado de Direito* (CETS n.º 225).

na investigação, de 2023), as Orientações Éticas da Comissão Europeia para educadores (2022) e as *Living Guidelines on the Responsible Use of Generative AI in Research* (2024).[16] Este acervo partilha uma orientação comum, a abordagem centrada no ser humano (*human-centred approach*), que o presente estudo adota e densifica no plano pedagógico. As universidades portuguesas estão, pois, juridicamente convocadas a definir políticas de uso da IA no ensino, e a elas se destina, em última análise, o modelo aqui proposto.

A Universidade de Lisboa deu, entretanto, um passo nesse sentido. Em outubro de 2026, a Reitoria divulgou à comunidade académica a atualização dos seus serviços digitais, com a disponibilização de ferramentas de IA a todos os estudantes e docentes e com orientações para a utilização responsável da IA generativa.[17] Duas notas desse documento merecem registo, pela sintonia com a tese que aqui se defende. A primeira é a justificação da medida nos termos da equidade: nas palavras do Reitor, Luís Ferreira, uma tecnologia com potencial para democratizar o conhecimento não pode tornar-se, ela própria, um novo fator de desigualdade.[18] É o quarto princípio ético da secção IV, formulado pela instituição. A segunda é a finalidade enunciada, dar à comunidade mais meios para aprender, criar, investigar e colaborar, com criatividade, pensamento crítico e responsabilidade, e o próprio portal institucional afirma que a IA é ferramenta de apoio e não substitui o conhecimento, o pensamento crítico ou a decisão humana, recomendando que se validem os resultados, se partilhe apenas a informação necessária e se respeitem as regras definidas pelos docentes nas avaliações.[19] São, uma a uma, as regras operativas que este estudo propõe. Não deixa de ser significativo que o presente Projeto, apresentado em junho de 2026, tenha antecipado estas orientações, e que a comunicação tenha sido dirigida à Faculdade de Direito por encargo do

16 UNESCO (2021). *Recommendation on the ethics of artificial intelligence*; Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO; Comissão Europeia (2022). *Orientações éticas sobre a utilização da inteligência artificial (IA) e de dados no ensino e na aprendizagem para os educadores*; e Comissão Europeia (2024). *Living guidelines on the responsible use of generative AI in research*.

17 Universidade de Lisboa, Reitoria. *Serviços digitais ULisboa: mais ferramentas, novas possibilidades*. Comunicação à comunidade académica, divulgada à Faculdade de Direito da Universidade de Lisboa pelo Secretariado dos Órgãos, por encargo do Diretor, em 14 de outubro de 2026. Portal em https://digital.ulisboa.pt.

18 Citação do Reitor da Universidade de Lisboa, Professor Doutor Luís Ferreira, na comunicação referida em [^ul1].

19 Universidade de Lisboa. *Serviços de inteligência artificial*. Disponível em https://digital.ulisboa.pt/info/servicos-de-inteligencia-artificial. A oferta inclui o *Copilot Chat* (Microsoft), o *Gemini*, o *NotebookLM*, o *Google Workspace*, os *Gems* e o *Workspace Studio* (Google) e a plataforma *IAedu* (FCT). O portal fixa, como orientações de uso responsável, que a IA é ferramenta de apoio e não substitui o conhecimento, o pensamento crítico ou a decisão humana, devendo validar-se os resultados, partilhar-se apenas a informação necessária e respeitar-se as regras definidas pelos docentes nas avaliações.

seu Diretor, o Professor Doutor Eduardo Vera-Cruz Pinto, que é também o orientador do Projeto. A convergência entre a proposta individual e a política institucional confirma a oportunidade do tema e reforça o seu interesse para a Faculdade.

IV. A ÉTICA COMO CRITÉRIO ESTRUTURANTE: CINCO PRINCÍPIOS

A ética não é, neste estudo, um capítulo acessório de considerações finais: é o critério estruturante de toda a proposta. A razão radica na natureza do ensino jurídico. Ensinar o Direito não é transferir informação: é formar a *prudentia*, a capacidade de julgar o caso, de ponderar o justo, de responder por decisões. Um instrumento que produz texto sem responsabilidade, sem verdade e sem juízo, porque um LLM não tem nenhuma das três, só pode entrar na formação de juristas subordinado a quem as tem: o professor e, em formação, o estudante. Invertida a hierarquia, o instrumento corrompe o fim. Importa recordar, com a crítica de Emily Bender e colegas aos papagaios estocásticos, o que os modelos de linguagem são e o que não são: não compreendem, não sabem, não decidem; produzem texto estatisticamente plausível.[20]

Daí que o estudo adote, como matriz ética operativa, cinco princípios, convergentes com o quadro unificado de Floridi e Cowls, com a Recomendação da UNESCO e com a Carta Ética da CEPEJ, reformulados para o contexto pedagógico.[21]

O primeiro é a *transparência*. Todo o uso de IA, pelo docente na preparação de materiais e pelo estudante nos trabalhos, é declarado. A declaração de uso de IA torna-se peça normal do trabalho académico, como a bibliografia. Transparência não é suspeição: é a condição que permite discutir o uso em vez de o esconder.

O segundo é a *autoria e integridade*. A IA pode assistir; não pode assinar. O trabalho avaliado exprime o juízo do estudante; a aula exprime o juízo do professor. A fraude não muda de natureza por mudar de instrumento, e os regulamentos de avaliação das faculdades fornecem a base disciplinar aplicável.

---

20 Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots. *FAccT '21*, pp. 610-623.

21 Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1); e Floridi, L. (2023). *The ethics of artificial intelligence*. Oxford University Press. No plano do problema do controlo, Russell, S. (2019). *Human compatible*. Viking; sobre os custos materiais e sociais, Crawford, K. (2021). *Atlas of AI*. Yale University Press.

O terceiro é a *equidade*. As ferramentas adotadas devem ser de acesso universal e tendencialmente gratuito, privilegiando os recursos abertos e os níveis gratuitos dos LLM, para que a IA não acrescente uma nova desigualdade socioeconómica entre estudantes. Um ensino que pressuponha subscrições pagas transforma a inovação em privilégio.

O quarto é a *proteção da pessoa e dos dados*. Nenhum dado pessoal de estudantes é confiado a sistemas de IA sem base jurídica e minimização (RGPD); são recusadas as aplicações proibidas ou de risco inaceitável, designadamente o reconhecimento de emoções em contexto educativo, vedado pelo artigo 5.º do Regulamento (UE) 2024/1689.

O quinto é a *responsabilidade e supervisão humana*. Toda a saída de IA usada em contexto letivo é verificada por um humano antes de produzir efeitos; nenhuma classificação, sanção ou juízo sobre pessoas é automatizado, por transposição, como princípio pedagógico, dos artigos 22.º do RGPD e 14.º do Regulamento IA. O ser humano permanece, na expressão consagrada, *in the loop* e, mais do que isso, *in command*.

Estes princípios não são declarações de intenção: traduzem-se, como se verá, em regras operativas verificáveis, e são o solo sobre o qual assenta o segundo fundamento da tese.

V. A DECISÃO HUMANA: DO *IUDEX* AO PROFESSOR

A insistência na decisão humana tem, neste estudo, um fundamento histórico-jurídico próprio, que constitui talvez o seu contributo mais original. A experiência jurídica romana, matéria do ensino do autor, mostra que o Direito nasceu e amadureceu como obra de pessoas concretas que respondiam pelas suas opiniões: o *prudens* que dava *responsa* com a sua *auctoritas*; o pretor que, no edicto, dizia em que termos daria a ação; o *iudex* que julgava vinculado à fórmula, mas livre no juízo sobre os factos.

Quando o poder normativo se concentrou e despersonalizou, da *lex* ao rescrito imperial, da *iuris prudentia* livre à Lei das Citações de 426, que mandava contar as opiniões de juristas mortos como quem conta votos, o Direito ganhou em certeza o que perdeu em vida. A Lei das Citações é, *avant la lettre*, o primeiro algoritmo de decisão jurídica da história ocidental: um mecanismo que substitui o juízo pela contagem, que resolve a divergência interpretativa por uma regra de maioria aplicada mecanicamente a autoridades predeterminadas. O seu estudo, que integra o programa de Direito Romano, é a melhor advertência pedagógica contra

a tentação de hoje: resolver a incerteza do juízo humano pela automação da resposta. A leitura da degradação da criação jurisprudencial perante a imposição legislativa, desenvolvida no magistério romanístico da FDUL, fornece precisamente esta chave.[22]

Esta leitura ganha densidade quando se percorre a longa duração do próprio ensino jurídico, que é, em si mesma, uma história de tecnologias intelectuais e de redistribuições de autoridade sobre o Direito. O *mos italicus* dos comentadores e o *mos gallicus* dos humanistas foram dois modos de ler e ensinar as fontes, cada um com a sua técnica; a imprensa fixou e difundiu os textos, alterando a relação do jurista com a autoridade; os Estatutos Pombalinos de 1772, com a Lei da Boa Razão no seu horizonte, foram eles próprios uma reforma metodológica e tecnológica do ensino do Direito, que subordinou a autoridade dos doutores à razão e às fontes.[23] De cada uma destas mutações a Universidade saiu transformada, mas não abdicou do seu ofício: integrou a nova técnica e reforçou, em vez de abandonar, os exercícios de juízo pessoal. A escrita não matou a *prudentia*, a imprensa não matou a Universidade, o código não matou a interpretação. A Inteligência Artificial generativa é a mais recente dessas tecnologias intelectuais, e o seu enquadramento correto é o que a Universidade sempre deu às anteriores.

Transposto para o ensino, o professor é, na sala de aula, o que o *prudens* era no *ius*: a instância pessoal de *auctoritas* que responde pelo que ensina. Esta leitura encontra apoio direto na investigação histórico-jurídica sobre a função jurisdicional e o estatuto do juiz na tradição portuguesa e ocidental, que mostra como a legitimidade da decisão sempre se ancorou na pessoa que a profere e por ela responde.[24] A IA pode ser o seu *scriba*, o seu *amanuensis*, e não deixa de ser eloquente que a principal ferramenta romanística digital tenha precisamente esse nome, mas nunca o seu substituto. E o estudante deve ser formado não como consumidor

22 Vera-Cruz Pinto, E. (2022). *Introdução ao direito romano: As questões fundamentais*. AAFDL. A leitura da relação entre criação jurisprudencial e imposição legislativa é aí desenvolvida como chave de compreensão da experiência jurídica romana.

23 Sobre a longa duração do ensino do Direito Romano em Portugal, Sampaio e Mello, G. de (1997). *Apontamentos para a história do ensino do direito romano em Portugal*. AAFDL; e, para o enquadramento da cultura jurídica europeia na longa duração, Hespanha, A. M. (2012). *Cultura jurídica europeia: Síntese de um milénio*. Almedina.

24 Barbas Homem, A. P. (2003). *Judex perfectus: Função jurisdicional e estatuto judicial em Portugal (1640-1820)*. Almedina; e Videira, S. Antas (Coord.) (2009). *O perfil do juiz na tradição ocidental*. Almedina. Ambos os estudos fundamentam a transposição da figura do juiz responsável para a figura do professor responsável perante a tentação da automação.

de respostas, mas como futuro decisor: alguém que um dia assinará pareceres, sentenças, contratos, leis, atos pelos quais responderá pessoalmente.

Da primazia da ética e da decisão humana decorrem, para o modelo pedagógico, cinco regras operativas que convém fixar com clareza, porque é nelas que a tese se torna verificável: primeira, a IA nunca avalia, pode gerar exercícios de autoavaliação formativa, mas a classificação é ato humano indelegável; segunda, a IA nunca é fonte, é instrumento de acesso a fontes, e toda a afirmação relevante se verifica na fonte primária ou na doutrina identificada; terceira, o uso de IA é sempre visível, declarado, discutido em aula, integrado na reflexão metodológica; quarta, há sempre uma via sem IA, nenhum estudante é obrigado a usar sistemas comerciais, garantindo-se alternativas; quinta, o tempo libertado pela IA é reinvestido no que só humanos fazem, discussão, leitura lenta das fontes, escrita pessoal, oralidade. Esta leitura inscreve-se, de resto, numa compreensão da juridicidade que recusa a redução do Direito a cálculo, e que tem no personalismo da Escola de Lisboa um dos seus solos mais firmes: transpor para a pedagogia a máxima com que Celso abre o *Digesto*, *ius est ars boni et aequi*, é afirmar que o Direito é arte, *prudentia*, e não mero cálculo, e que o seu ensino também o é.[25]

VI. AS FERRAMENTAS DE IA APLICÁVEIS: INVENTÁRIO POR FUNÇÃO

Inventariam-se agora as ferramentas com utilidade demonstrável para o ensino, agrupadas por função pedagógica e não por produto, porque os produtos envelhecem e as funções permanecem. Exemplifica-se a aplicação aos programas de Direito Romano e de História do Direito Português, mas o critério de seleção, acessibilidade, custo, proteção de dados e fiabilidade, é extensível a qualquer disciplina.

Convém começar pelas plataformas a que o estudante da Universidade de Lisboa tem acesso institucional, porque é delas que se deve partir. Com a atualização do *Google Workspace* para a edição *Education Plus*, divulgada à comunidade académica em outubro de 2026, a ULisboa colocou à disposição de todos os estudantes, sem custo acrescido, o *Gemini* e o *NotebookLM* para analisar documentos, resumir informação e aprofundar temas, além de

25 Vera-Cruz Pinto, E. (2015). *O futuro da justiça*. Nova Vega. A recusa da redução do Direito a cálculo é aí desenvolvida como eixo de uma compreensão personalista da juridicidade.

funcionalidades de IA na criação de apresentações, documentos e vídeos e no *Google Meet*.[26] A par destas, a *Microsoft* disponibiliza o *Copilot Chat* no ecossistema *Microsoft 365*, e a Fundação para a Ciência e a Tecnologia mantém a plataforma IAedu, com acesso a vários modelos e possibilidade de criar assistentes temáticos.[27] Esta oferta institucional tem uma consequência direta para o princípio da equidade: o docente pode desenhar exercícios que pressupõem o uso de IA sabendo que nenhum estudante fica de fora por não poder pagar uma subscrição. As descrições que se seguem partem, por isso, destas plataformas, e indicam, em cada caso, o que o estudante faz na prática.

a) ***Modelos de linguagem de grande escala***

O *ChatGPT* (OpenAI), o *Claude* (Anthropic), o *Gemini* (Google) e o *Copilot* (Microsoft) são os assistentes generalistas de referência, e os dois últimos estão, como se viu, ao alcance institucional do estudante da ULisboa. Na prática, o estudante usa-os de quatro modos.

No primeiro, a tutoria socrática, não pede ao modelo a resposta, mas a crítica da sua resposta. Escreve, por palavras próprias, a distinção entre *auctoritas* e *potestas*; cola o texto no assistente com a instrução de apontar lacunas e contra-argumentos; lê a crítica e decide o que aceita e o que rejeita, justificando. O modelo deixa de ser oráculo e passa a ser parceiro de treino.

No segundo, a simulação de controvérsia, o estudante pede ao assistente que assuma a posição de uma das escolas, por exemplo a dos sabinianos sobre a *specificatio*, e argumenta contra ela, obrigando-se a construir o contraditório; depois confronta o diálogo com as fontes e verifica qual das posições o Direito acolheu.

No terceiro, o apoio ao latim jurídico, pede uma tradução de trabalho de um fragmento do *Digesto* e compara-a, linha a linha, com o texto da edição crítica acedido no *Amanuensis*,

26 Universidade de Lisboa, Reitoria. *Serviços digitais ULisboa: mais ferramentas, novas possibilidades*. Comunicação à comunidade académica, divulgada à Faculdade de Direito da Universidade de Lisboa pelo Secretariado dos Órgãos, por encargo do Diretor, em 14 de outubro de 2026. Portal em https://digital.ulisboa.pt.

27 Universidade de Lisboa. *Serviços de inteligência artificial*. Disponível em https://digital.ulisboa.pt/info/servicos-de-inteligencia-artificial. A oferta inclui o *Copilot Chat* (Microsoft), o *Gemini*, o *NotebookLM*, o *Google Workspace*, os *Gems* e o *Workspace Studio* (Google) e a plataforma *IAedu* (FCT). O portal fixa, como orientações de uso responsável, que a IA é ferramenta de apoio e não substitui o conhecimento, o pensamento crítico ou a decisão humana, devendo validar-se os resultados, partilhar-se apenas a informação necessária e respeitar-se as regras definidas pelos docentes nas avaliações.

anotando onde a tradução automática simplificou ou errou um termo técnico. A regra de ouro, que o docente repete, é que a tradução da máquina é hipótese, nunca resultado.

No quarto, a geração de materiais de estudo, pede ao modelo um conjunto de perguntas de autoavaliação sobre a Lei das XII Tábuas ou um resumo de um bloco da matéria, que usa para testar o que já sabe, e que o docente validou previamente. O exemplo pedagógico decisivo é, porém, o do erro: porque os modelos alucinam fontes e cometem anacronismos, o docente pede deliberadamente uma resposta que sabe que virá com defeitos, e transforma a sua correção em exercício, assunto a que a secção VII dá tratamento próprio.

*b) **Ambientes de estudo ancorados em corpora controlados***

O *NotebookLM* (Google), também ele ao alcance institucional do estudante da ULisboa, permite criar cadernos de estudo alimentados exclusivamente por documentos carregados pelo utilizador, com citação localizada das passagens, o que reduz drasticamente a alucinação. Na prática, o docente prepara um caderno com o manual adotado e uma coletânea de fontes, partilha-o com a turma, e o estudante passa a interrogá-lo por escrito: pergunta em que título se regula determinada matéria e recebe a resposta com a indicação da página ou do fragmento de onde ela foi retirada, o que lhe permite ir imediatamente verificar. Um caderno assim construído é um tutor que responde apenas pelo *corpus* autorizado, modelo de *retrieval-augmented generation* (RAG) que o estudo defende como arquitetura preferencial para o ensino, por manter o estudante preso à fonte. Os *Claude Projects* e os *GPTs* personalizados permitem o mesmo desenho, e a plataforma *IAedu* da FCT acrescenta a possibilidade de criar assistentes temáticos da unidade curricular. Configurar um destes assistentes é, em si mesmo, ato pedagógico do professor, que escolhe o que entra no *corpus* e o que dele fica de fora, e por isso confirmação prática da tese da decisão humana.

c) Pesquisa académica assistida

O *Elicit*, o *Consensus*, o *Scite*, o *Semantic Scholar* e o *Research Rabbit* assistem à revisão de literatura. Na preparação de um relatório de aula prática, o estudante formula a sua pergunta de investigação no *Elicit* e obtém uma lista de artigos com a síntese das respetivas conclusões; usa o *Research Rabbit* para, a partir de um texto que já conhece, descobrir os trabalhos que o citam e os que ele cita, construindo o mapa da discussão; e, no

Scite, verifica se uma afirmação é apoiada ou contestada pela literatura através das indicações de citação. Em todos, a tarefa seguinte é a mesma e é humana: ler os textos que interessam e decidir quais entram no relatório, porque a ferramenta localiza, mas não lê pelo estudante. O *Perplexity* e os modos de pesquisa profunda dos assistentes principais agregam fontes da web com ligação direta, o que facilita a verificação, mas não a dispensa.

d) Humanidades digitais e fontes histórico-jurídicas

O *Amanuensis*, construído por Peter Riedlberger e Günther Rosenbaum sobre a base *ROMTEXT* da Universidade de Linz, é a ferramenta nuclear para o Direito Romano: pesquisa instantânea em todo o *Corpus Iuris Civilis*, no *Codex Theodosianus*, em Gaio e nas leis germânicas, gratuita e multiplataforma, o que significa que qualquer estudante pode trazer no telemóvel, para a aula prática, a totalidade pesquisável do *Digesto*.[28] O *Transkribus*, gerido pela cooperativa READ-COOP, aplica reconhecimento automático de escrita manuscrita (*handwritten text recognition*) a manuscritos, permitindo que, num seminário avançado, os estudantes transcrevam fólios de chancelaria régia ou de praxistas portugueses.[29] O portal *Ius Lusitaniae* (Universidade Nova de Lisboa), os acervos da Torre do Tombo e da Biblioteca Nacional de Portugal fornecem os *corpora*; ferramentas de análise textual, como o *Voyant Tools* e o *Classical Language Toolkit* para latim, completam o instrumentário. Os tradutores neuronais assistem o trabalho sobre a grande romanística europeia, democratizando o acesso à bibliografia estrangeira.

e) Acessibilidade, oralidade e organização do estudo

Sistemas de transcrição automática como o *Whisper* (OpenAI) respondem, com consentimento, a necessidades educativas especiais; sintetizadores de voz tornam as fontes audíveis; os LLM geram glossários, cronologias e mapas conceptuais a partir dos materiais da cadeira, sempre revistos pelo docente. As ferramentas de deteção de texto gerado por IA são deliberadamente excluídas do modelo: a sua fiabilidade é insuficiente, com taxas de

28 Riedlberger, P., & Rosenbaum, G. (2014 e seguintes). *Amanuensis* [software, base ROMTEXT, Universidade de Linz]. Disponível em https://www.riedlberger.de/amanuensis.
29 READ-COOP. *Transkribus*. Disponível em https://www.transkribus.org. Veja-se ainda Kahle, P., Colutto, S., Hackl, G., & Mühlberger, G. (2017). Transkribus: A service platform for transcription. *ICDAR*, pp. 19-24.

falsos positivos incompatíveis com a presunção de inocência, e o seu uso punitivo automatizado violaria o princípio da decisão humana.

f) Síntese e limitações

O inventário confirma a estrutura da tese: as ferramentas valem o desenho humano do seu uso. O mesmo LLM pode ser máquina de copiar ou ginásio de espírito crítico; a diferença está no professor. O inventário não ficaria, porém, completo sem as limitações, que não infirmam a tese, antes a integram, por serem precisamente o tipo de juízo prudencial que só a decisão humana pode fazer. A primeira é a dependência de fornecedores privados, maioritariamente extraeuropeus, com modelos de negócio e políticas de dados que a Universidade não controla, razão para privilegiar ferramentas abertas e *corpora* próprios. A segunda é linguística: os LLM têm desempenho inferior em português europeu face ao inglês, e inferior ainda em latim jurídico e em português clássico, o que constitui, em si, um resultado científico relevante. A terceira são os custos, que a equidade manda neutralizar. A quarta é a pegada ambiental e os custos sociais da IA, que devem ser tematizados com os estudantes como parte da literacia crítica exigida pelo artigo 4.º do Regulamento (UE) 2024/1689.[30] A quinta é a proteção de dados e a confidencialidade, tratada autonomamente adiante.

VII. DEMONSTRAÇÃO PRÁTICA: QUATRO GUIÕES DE AULA

Para demonstrar que a tese não é programática, mas operativa, apresentam-se quatro guiões de aula completos, com objetivo, sequência, ponto de verificação crítica e produto avaliável, sempre sob as regras éticas da secção IV. Estes guiões foram já experimentados, em versão preliminar, nas aulas práticas do autor.

Demonstração 1: Direito Romano, auctoritas e potestas (tutoria socrática controlada)

O objetivo é levar o estudante a distinguir *auctoritas* e *potestas* e a reconhecer a tendência dos LLM para reduzir o Direito à lei e ao poder estatal. A sequência tem quatro

30 Crawford, K. (2021). *Atlas of AI*. Yale University Press. A tematização dos custos ambientais e sociais integra a literacia crítica exigida pelo artigo 4.º do Regulamento (UE) 2024/1689.

passos: cada estudante redige, sem IA, a sua explicação da distinção; submete-a a um LLM e pede crítica; confronta a crítica com a lição do manual; e regista por escrito onde o modelo acertou, onde errou e porquê. O ponto de verificação crítica é previsível: tipicamente o modelo reproduz uma visão legalista e estatizante, exatamente a que o programa ensina a criticar, e o erro da máquina torna-se a porta de entrada da tese do curso. O produto avaliável é uma ficha de uma página com a explicação própria, a crítica do modelo e a correção fundamentada.

Demonstração 2: Direito Romano, a alucinação de fontes como exercício de crítica

O objetivo é ensinar, em simultâneo, a epistemologia da ferramenta e a crítica de fontes. A sequência: o docente pede a um LLM uma resposta sobre os *senatusconsultos* como fonte de *ius*, com indicação de passagens do *Digesto*; distribui a resposta à turma; os estudantes verificam cada citação no *Amanuensis* e na edição crítica; anotam as citações inexistentes ou deslocadas, isto é, alucinadas, e as afirmações anacrónicas; e reescrevem a resposta corrigida, com fontes verdadeiras. O ponto de verificação crítica é a confrontação de cada *inscriptio* com a fonte real, pela própria mão do estudante. O produto avaliável é a resposta corrigida e anotada, com identificação expressa de cada erro. Este exercício materializa a *interpretatio prudentium* aplicada à IA: o jurista que verifica o texto de autoridade duvidosa, que é o que a literatura documenta como a fragilidade mais característica dos LLM em matéria jurídica.[31]

Demonstração 3: História do Direito Português, as três Ordenações (RAG e tabela comparativa)

O objetivo é comparar o tratamento de um mesmo instituto nas Ordenações Afonsinas, Manuelinas e Filipinas, com apoio de um ambiente ancorado. A sequência: o docente configura um caderno *NotebookLM* alimentado apenas com os textos das Ordenações, acedidos via *Ius Lusitaniae*, e com uma coletânea de fontes; os estudantes interrogam o *corpus* com citação localizada; a IA gera uma primeira tabela comparativa; e os estudantes corrigem-na sobre os textos. O ponto de verificação crítica é a correção da tabela sobre as

31 Dahl, M., Magesh, V., Suzgun, M., & Ho, D. E. (2024). Large legal fictions: Profiling legal hallucinations in large language models. *Journal of Legal Analysis*, 16(1), pp. 64-93.

fontes, aprende-se mais ao corrigir do que ao compilar. O produto avaliável é a tabela comparativa final validada, com indicação das passagens.

Demonstração 4: História do Direito Português, escolas historiográficas (deteção de homogeneização)

O objetivo é mostrar que os LLM tendem a homogeneizar perspetivas historiográficas distintas, e ensinar a distingui-las. A sequência: pede-se ao modelo que exponha a codificação oitocentista segundo diferentes autores; os estudantes, com os livros na mesa, verificam se o modelo distingue realmente as escolas ou as funde; e reconstroem as diferenças reais com apoio da bibliografia. O ponto de verificação crítica é a comparação entre o que o modelo afirma e o que os autores efetivamente sustentam, e a descoberta da homogeneização vale uma aula inteira sobre o que é uma escola historiográfica. O produto avaliável é um quadro das perspetivas, com citações verdadeiras e identificação do erro de homogeneização do modelo.

Dois instrumentos próprios de integração: o Gabinete do Professor e o jogo da semana

Os quatro guiões anteriores dizem respeito ao uso da IA no interior de cada aula. A integração, porém, não se esgota na sessão letiva: exige um ambiente que sustente, ao longo do semestre, a relação entre o docente, as turmas e os estudantes, e que torne a decisão humana sobre a avaliação não só possível, mas documentada. É essa a função de dois instrumentos próprios, concebidos e postos em prática pelo autor, que materializam, no plano institucional e no plano da avaliação contínua, a mesma tese.

O primeiro é o Gabinete do Professor, uma aplicação web desenvolvida pelo autor e em produção, que reúne num painel único a gestão das disciplinas, das turmas e dos estudantes, uma biblioteca jurídica e de materiais multimédia, o espaço de projetos e uma gestão de acessos por perfis (administrador, colaborador e estudante), com funcionalidades de exportação e auditoria orientadas ao cumprimento do RGPD. A relevância da plataforma para a tese é dupla. Por um lado, constitui o ambiente controlado que os princípios de proteção de dados e de equidade exigem: em vez de confiar listas de alunos, classificações e materiais a sistemas comerciais extraeuropeus, o docente dispõe de um espaço próprio, cujos fluxos de

dados conhece e governa, e onde os assistentes de estudo ancorados em *corpus* autorizado (secção VI) podem ser disponibilizados sem expor dados pessoais a terceiros. Por outro lado, a própria construção da plataforma é a demonstração mais literal da tese da decisão humana: o professor não é utilizador passivo de ferramentas alheias, mas autor do ambiente em que a tecnologia entra, e é essa autoria que lhe permite manter-se *in command*.

O segundo instrumento é o jogo da semana, um dispositivo de avaliação contínua gamificada. Cada estudante dispõe de um código QR que contém a sua identificação, e a leitura desse código, ao longo da atividade semanal, regista a participação de cada um nas tarefas propostas (resolução de casos, crítica de fontes, correção de respostas geradas por IA), alimentando o registo de avaliação contínua. O valor pedagógico do dispositivo é triplo. Primeiro, resolve um problema real das turmas numerosas, dar visibilidade individual à participação de cada estudante sem sacrificar o ritmo da aula, o que serve diretamente o objetivo de individualizar o apoio num contexto de massificação. Segundo, torna a avaliação contínua um registo de processo, e não apenas de produto, em sintonia com o desenho de avaliação defendido na secção VIII: o que fica registado é o percurso de participação, verificável e discutível. Terceiro, e decisivamente, o dispositivo ilustra o limite ético que a tese fixa: o código QR e o sistema recolhem e organizam dados de participação, mas não classificam; a leitura do código é captura, não juízo. A classificação permanece ato humano fundamentado, indelegável, exercido pelo professor sobre a informação assim reunida. O tratamento de dados pessoais que o dispositivo implica (os nomes dos estudantes) obedece, pela sua inserção num ambiente institucional controlado como o Gabinete, aos princípios de minimização e de licitude do RGPD, e não a sistemas externos de verificação automática de identidade, que o modelo recusa.

Os dois instrumentos convergem no mesmo ponto dos quatro guiões: a tecnologia amplia o que o professor pode fazer, mas é o professor que a desenha, a governa e, no momento decisivo da avaliação, decide.

O que estes exemplos demonstram

Três conclusões emergem da demonstração. A primeira: nenhum dos usos descritos dispensa o professor, todos exigem desenho prévio, validação de conteúdos e condução crítica, pelo que a IA aumenta, e não reduz, a exigência sobre o docente. A segunda: todos

os usos convergem no mesmo gesto pedagógico, devolver o estudante às fontes, com mais meios e mais motivos para as ler. A terceira: as disciplinas histórico-jurídicas, longe de serem as mais ameaçadas pela IA, são as mais bem armadas para a domesticar, porque o seu método, a crítica de fontes, é exatamente o método de que a era da IA generativa precisa. Formar juristas que verificam é a contribuição própria da História do Direito para a Universidade do século XXI, e é esta a competência que, como adiante se sustenta, deve irradiar para todo o currículo.

VIII. AVALIAÇÃO, INTEGRIDADE ACADÉMICA E DESENHO DE PROVAS

A avaliação é o ponto nevrálgico, e por isso merece desenvolvimento autónomo. A IA generativa tornou inseguro, como instrumento exclusivo de avaliação, o trabalho escrito não presencial sem defesa: qualquer enunciado genérico obtém de um LLM resposta de qualidade média em segundos. A reação errada seria a vigilância tecnológica. Os detetores de texto gerado por IA têm taxas de falsos positivos incompatíveis com o princípio da presunção de inocência e penalizam desproporcionadamente estudantes não nativos; e os sistemas de *proctoring* algorítmico de exames caem na qualificação de risco elevado do Anexo III do Regulamento (UE) 2024/1689, além de instaurarem uma pedagogia da suspeita contrária à relação de confiança que o ensino exige.[32]

A reação correta, que o modelo adota e o protótipo testa, é o desenho de avaliação. São cinco as linhas de ação. Primeira, deslocar o peso para o presencial e para o oral, recuperando a tradição da arguição que as disciplinas histórico-jurídicas nunca abandonaram. Segunda, nos trabalhos não presenciais, avaliar o processo e não apenas o produto, exigindo o registo do percurso (fontes consultadas, interações com IA declaradas, versões) e a defesa oral do trabalho. Terceira, construir enunciados resistentes à IA por serem ancorados no concreto da aula. Quarta, usar a própria IA como objeto de avaliação, pedindo ao estudante que identifique e corrija os erros de uma resposta gerada por um modelo, exercício que nenhuma delegação na máquina resolve. Quinta, manter sempre a classificação como ato humano fundamentado, suscetível de revisão nos termos dos regulamentos de avaliação aplicáveis. A

32 Regulamento (UE) 2024/1689, Anexo III, ponto 3, e artigo 5.º. Sobre a insuficiência dos detetores de texto gerado por IA, veja-se a reflexão crítica em Holmes, W., & Porayska-Pomsta, K. (Eds.) (2023). *The ethics of artificial intelligence in education*. Routledge.

integridade académica defende-se, em suma, tornando o juízo humano inevitável, e não tentando, em vão, tornar a máquina indetetável ou detetável.

Instrumento central desta arquitetura é a declaração de utilização de ferramentas de IA, que acompanha os trabalhos escritos. A declaração cumpre três funções pedagógicas: normaliza a transparência, porque ninguém é punido por declarar, ao passo que a omissão, essa sim, é fraude; educa para a verificação, já que a afirmação de que todas as passagens assistidas por IA foram confirmadas nas fontes é, em si, um programa metodológico; e converte o uso da IA em objeto de reflexão avaliável, ao valorizar a síntese das interações, dos erros detetados e das correções feitas. É, em miniatura, toda a tese: a ferramenta entra, mas quem responde é a pessoa.

IX. SEGURANÇA DA INFORMAÇÃO, FONTES E PROPRIEDADE INTELECTUAL

O uso da IA no ensino convoca três exigências que o docente e o estudante não podem ignorar, e que o modelo converte em regras operativas. Não são acessórios técnicos: são dimensões do mesmo dever de prudência que estrutura toda a proposta.

A primeira é a segurança da informação e a proteção de dados. A introdução de sistemas de IA implica fluxos de dados que devem ser desenhados à luz do RGPD. As regras são de minimização (não se introduzem em *prompts* dados pessoais de estudantes nem material confidencial), de base jurídica e transparência (qualquer tratamento de dados por ferramentas institucionais tem base identificada e é comunicado), de preferência por ambientes controlados (privilegiam-se soluções que não usem os dados introduzidos para treino de modelos, e o processamento local sempre que possível), de consentimento informado e revogável para a componente empírica, e de segregação (o material não publicado de terceiros nunca é carregado em sistemas externos sem autorização escrita).[33]

A segunda é a certeza e verificação das fontes. O risco mais característico dos LLM, em matéria jurídica, é a alucinação: a invenção de fragmentos do *Digesto*, de títulos das Ordenações, de acórdãos ou de doutrina inexistente, apresentados com aparência de

33 Regulamento (UE) 2016/679 (RGPD), em especial os artigos 5.º (princípios, incluindo minimização e limitação da finalidade), 6.º (licitude), 9.º (categorias especiais) e 22.º (decisões automatizadas).

autoridade. A literatura documenta-o em escala preocupante.[34] O modelo faz desta fragilidade um instrumento pedagógico, mas fixa, antes, uma regra inegociável: nenhuma afirmação relevante assistida por IA entra num trabalho ou numa aula sem verificação na fonte primária ou na doutrina identificada e consultada. A IA é instrumento de acesso a fontes; nunca é, ela própria, fonte. Esta regra é, a um tempo, exigência de rigor científico e treino da competência profissional central do jurista: responder apenas pelo que pode verificar.

A terceira é a propriedade intelectual e os direitos de autor, que se coloca em dois planos. No plano dos *inputs*, não se introduzem em sistemas de IA obras protegidas de terceiros para além do uso lícito, nem se faz carregamento de materiais cujos termos de licença o proíbam. No plano dos *outputs*, o texto gerado por IA não tem, em regra, autor para efeitos de direito de autor, que pressupõe uma criação intelectual de pessoa humana, nos termos dos artigos 1.º e 2.º do Código do Direito de Autor e dos Direitos Conexos; apresentar como obra própria um texto gerado pela máquina é, por isso, a um tempo, fraude académica e usurpação imprópria. O estudante é ensinado a tratar o *output* da IA como matéria-prima a verificar e reelaborar, sempre sob a sua autoria e responsabilidade, e a declará-lo.[35]

Os três planos convergem num único princípio, que é o do estudo inteiro: a ferramenta entra, mas quem responde, pela segurança dos dados, pela verdade das fontes e pela autoria do texto, é sempre a pessoa.

X. DA HISTÓRIA DO DIREITO A TODAS AS DISCIPLINAS JURÍDICAS: O PROJETO DE PÓS-DOUTORAMENTO

O que até aqui se demonstrou a partir do Direito Romano e da História do Direito Português não vale apenas para elas. Foi deliberada a escolha das disciplinas de fontes como laboratório, pela nitidez com que nelas o erro do modelo se expõe e se corrige; mas o critério que delas se extrai, a primazia da ética e da decisão humana, e a conversão da falibilidade da IA em exercício de verificação, é um critério geral do ensino jurídico. É este o horizonte do

34 Dahl, M., Magesh, V., Suzgun, M., & Ho, D. E. (2024). Large legal fictions: Profiling legal hallucinations in large language models. *Journal of Legal Analysis*, 16(1), pp. 64-93.

35 Código do Direito de Autor e dos Direitos Conexos, artigos 1.º e 2.º, sobre a exigência de criação intelectual de pessoa humana como pressuposto da tutela autoral.

Projeto de Pós-Doutoramento em que o presente estudo se integra: desenvolver, para a formação dos juristas no seu todo, e não apenas para a História do Direito ou para as disciplinas afins, um modelo pedagógico de uso da IA eticamente fundado e centrado na decisão humana.

A transposição faz-se por analogia funcional. Onde a História do Direito confronta o *output* do modelo com o *Digesto* ou com as Ordenações, o Direito Civil confronta-o com o texto vigente do Código Civil e com a doutrina obrigacionista de referência; o Direito Penal, com o princípio da legalidade e com a jurisprudência dos tribunais superiores; o Direito Administrativo, com o Código do Procedimento Administrativo e com a jurisprudência administrativa; o Direito da União Europeia, com os Tratados, o direito derivado e a jurisprudência do Tribunal de Justiça. Em todas, o gesto pedagógico é o mesmo: a IA gera uma primeira resposta plausível, e o estudante aprende a disciplina ao verificá-la, corrigi-la e fundamentá-la contra a fonte autorizada. A alucinação de jurisprudência, que nas disciplinas dogmáticas é um risco profissional gravíssimo, documentado já em casos reais de peças forenses com acórdãos inventados, torna-se, em contexto letivo controlado, o melhor exercício de pesquisa e citação jurídica que se pode conceber. A crítica de fontes, método nativo da História do Direito, revela-se assim competência nuclear de toda a formação jurídica na era da IA.

A analogia não é abstrata; traduz-se em guiões transponíveis. No Direito Civil, o docente pede ao modelo um parecer sobre a responsabilidade civil por facto ilícito e manda os estudantes confrontá-lo com o artigo 483.º e seguintes do Código Civil e com a doutrina obrigacionista de referência, identificando onde o modelo confunde pressupostos, cita jurisprudência inexistente ou importa soluções de outros ordenamentos; o produto avaliável é o parecer corrigido, com indicação precisa de cada norma e de cada acórdão verificado. No Direito Penal, a resistência à IA é maior precisamente onde a disciplina é mais exigente: o princípio da legalidade e a proibição da analogia *in malam partem* tornam a verificação da fonte não um exercício académico, mas uma exigência constitucional, e o estudante que corrige a resposta gerada aprende, ao mesmo tempo, a matéria e o fundamento garantístico da sua verificação. No Direito Administrativo e no Direito da União Europeia, a tarefa é confrontar o *output* com o Código do Procedimento Administrativo, com os Tratados e com a jurisprudência dos tribunais competentes, incluindo a do Tribunal de Justiça. Em todos

estes casos se repete o gesto das disciplinas histórico-jurídicas: a alucinação de jurisprudência, que na prática forense é um risco gravíssimo, já documentado em peças reais com acórdãos inventados, converte-se, em contexto letivo controlado, no melhor exercício de pesquisa e citação que se pode conceber. Não se trata de uma extrapolação teórica. O autor aplica já este modelo nas disciplinas dogmáticas que rege na Universidade Europeia de Lisboa, designadamente nos Direitos Reais e no Direito da Família, onde os mesmos princípios, verificação obrigatória na lei e na jurisprudência, declaração do uso e classificação como ato humano, se revelam tão operativos quanto no Direito Romano.

Três razões sustentam a generalização. A primeira é de método: o Direito, em qualquer dos seus ramos, é uma disciplina de fontes e de juízo, e a IA interpela-o sempre nos mesmos dois pontos, o acesso à fonte, que ela amplia, e o juízo sobre a fonte, que ela não pode substituir. A segunda é normativa: o quadro europeu e português aplicável ao uso da IA no ensino (secção III) não distingue entre disciplinas, e a literacia em IA que o artigo 4.º do Regulamento (UE) 2024/1689 impõe é devida a todos os docentes e estudantes. A terceira é institucional: as faculdades de Direito precisam de uma política de uso da IA que seja uma só, coerente, e não um mosaico de práticas disciplinares avulsas; e essa política ganha em firmeza se radicar no critério que a tradição histórico-jurídica melhor guarda, o da responsabilidade pessoal pela decisão.

O Projeto prevê, por isso, três produtos de vocação transversal. O primeiro é um modelo pedagógico documentado e testado, com guiões, exercícios e critérios de avaliação, concebido para ser adaptado por qualquer docente, de qualquer disciplina jurídica. O segundo é uma bateria de avaliação das ferramentas, que mede o desempenho dos modelos não só em fontes históricas, mas também em legislação e jurisprudência vigentes, incluindo a medição específica do desempenho em português europeu e em latim jurídico. O terceiro é uma proposta de princípios para o uso da IA no ensino, suscetível de ser discutida nos órgãos competentes da Faculdade e de servir de base a uma política institucional. É no quadro deste programa, e como contributo para a renovação da pedagogia jurídica que ele persegue, que o modelo aqui exposto concorre ao Prémio de Inovação Pedagógica de 2026 da Universidade de Lisboa.

XI. CONCLUSÕES

1. A difusão dos modelos de linguagem de grande escala não é uma hipótese de futuro, mas o quotidiano das faculdades de Direito, e impõe uma resposta que não seja nem a proibição ineficaz nem a adoção acrítica, mas a integração crítica, eticamente fundada e pedagogicamente orientada.
2. O ensino do Direito será mais bem desenvolvido com a utilização da Inteligência Artificial, desde que essa utilização permaneça subordinada a dois fatores: a ética, como critério estruturante, e a decisão humana, como centro irredutível. A IA não substitui o professor; potencia-o. Não decide; auxilia a decidir. Não interpreta as fontes; multiplica o acesso a elas.
3. As disciplinas histórico-jurídicas, o Direito Romano e a História do Direito Português, constituem o laboratório ideal desta reflexão, por serem disciplinas de fontes, de contexto e de juízo, e porque a sua propensão a expor a alucinação e o anacronismo dos modelos as torna, não as mais ameaçadas pela IA, mas as mais bem armadas para a domesticar.
4. A ética do uso traduz-se em cinco princípios operativos e verificáveis: transparência, autoria e integridade, equidade, proteção da pessoa e dos dados, e responsabilidade com supervisão humana.
5. A decisão humana encontra fundamento histórico-jurídico próprio: o professor é, na sala de aula, o que o *prudens* era no *ius*, a instância pessoal de *auctoritas* que responde pelo que ensina; e a Lei das Citações de 426, primeiro algoritmo de decisão jurídica da história ocidental, é a melhor advertência contra a substituição do juízo pela contagem.
6. O fator decisivo do sucesso pedagógico não é a ferramenta, mas o desenho humano do seu uso. Os quatro guiões de aula apresentados demonstram que a IA, bem desenhada, aumenta a exigência sobre o docente e devolve o estudante às fontes.
7. A integridade académica defende-se pelo desenho de avaliação que torna o juízo humano inevitável, com deslocação para o oral e o presencial, avaliação do processo, uso da IA como objeto de avaliação e exclusão dos detetores automáticos e do *proctoring* algorítmico, e não pela vigilância tecnológica.

8. O critério extraído das disciplinas de fontes é um critério geral do ensino jurídico: por analogia funcional, o que a História do Direito faz com o *Digesto* e as Ordenações, cada disciplina dogmática faz com a lei vigente, a jurisprudência e a doutrina. A crítica de fontes, método nativo da História do Direito, é a competência nuclear de toda a formação jurídica na era da IA.
9. É este o horizonte do Projeto de Pós-Doutoramento em curso, e é como contributo para a renovação da pedagogia jurídica que o modelo aqui exposto concorre ao Prémio de Inovação Pedagógica de 2026 da Universidade de Lisboa. A História do Direito ensina que nenhuma técnica é neutra e nenhuma é fatal: a escrita não matou a prudência, a imprensa não matou a universidade, o código não matou a interpretação, mas cada mutação exigiu dos juristas uma reapropriação crítica. A Inteligência Artificial é a mutação do nosso tempo, e o seu enquadramento correto é o que a Universidade sempre deu às anteriores: com conhecimento, com ética e com a serena confiança de que o Direito é, e continuará a ser, obra de pessoas, *hominum causa omne ius constitutum sit* (D. 1, 5, 2).

REFERÊNCIAS

LEGISLAÇÃO E INSTRUMENTOS NORMATIVOS

CARTA DOS DIREITOS FUNDAMENTAIS DA UNIÃO EUROPEIA (2016). *Jornal Oficial da União Europeia*, C 202.

CÓDIGO DO DIREITO DE AUTOR E DOS DIREITOS CONEXOS (com alterações). *Diário da República*.

COMISSÃO EUROPEIA (2022). *Orientações éticas sobre a utilização da inteligência artificial (IA) e de dados no ensino e na aprendizagem para os educadores*. Serviço das Publicações da União Europeia.

COMISSÃO EUROPEIA (2024). *Living guidelines on the responsible use of generative AI in research*. European Research Area Forum.

COMISSÃO EUROPEIA PARA A EFICIÊNCIA DA JUSTIÇA (CEPEJ) (2018). *European ethical charter on the use of artificial intelligence in judicial systems and their environment*. Conselho da Europa.

CONSELHO DA EUROPA (2024). *Convenção-Quadro do Conselho da Europa sobre Inteligência Artificial e Direitos Humanos, Democracia e Estado de Direito* (CETS n.º 225).

CONSTITUIÇÃO DA REPÚBLICA PORTUGUESA (1976, com revisões). *Diário da República*.

DESPACHO N.º 8672/2021, de 1 de setembro. Regulamento sobre Programas de Pós-Doutoramento da Faculdade de Direito da Universidade de Lisboa. *Diário da República*, 2.ª série.

LEI N.º 27/2021, de 17 de maio. Carta Portuguesa de Direitos Humanos na Era Digital. *Diário da República*, 1.ª série, n.º 95.

REGULAMENTO (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016 (Regulamento Geral sobre a Proteção de Dados). *Jornal Oficial da União Europeia*, L 119.

REGULAMENTO (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024 (Regulamento Inteligência Artificial). *Jornal Oficial da União Europeia*, L, 2024/1689.

REPÚBLICA PORTUGUESA / INCoDe.2030 (2019). *AI Portugal 2030: Estratégia Nacional de Inteligência Artificial*. FCT.

UNESCO (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO.

DOUTRINA

ALETRAS, N., Tsarapatsanis, D., Preoţiuc-Pietro, D., & Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights. *PeerJ Computer Science*, 2, e93.

ASHLEY, K. D. (2017). *Artificial intelligence and legal analytics*. Cambridge University Press.

BARBAS HOMEM, A. P. (2001). *A lei da liberdade, I: Introdução histórica ao pensamento jurídico. Épocas medieval e moderna*. Principia.

BARBAS HOMEM, A. P. (2003). *Judex perfectus: Função jurisdicional e estatuto judicial em Portugal (1640-1820)*. Almedina.

BARBAS HOMEM, A. P. (2006). *O espírito das instituições: Um estudo de história do Estado*. Almedina.

BENDER, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots. *FAccT '21*, pp. 610-623.

BRIGAS, M. A. (2017). *As relações de poder na construção do direito da família português (1750-1910)*. AAFDL.

CARIDADE DE FREITAS, P. (2015). *História do direito internacional público (séculos XVI-XVIII)*. Quid Juris.

CHOI, J. H., Hickman, K. E., Monahan, A. B., & Schwarcz, D. (2022). ChatGPT goes to law school. *Journal of Legal Education*, 71(3), pp. 387-400.

COECKELBERGH, M. (2020). *AI ethics*. MIT Press.

CRAWFORD, K. (2021). *Atlas of AI*. Yale University Press.

DAHL, M., Magesh, V., Suzgun, M., & Ho, D. E. (2024). Large legal fictions: Profiling legal hallucinations in large language models. *Journal of Legal Analysis*, 16(1), pp. 64-93.

FLORIDI, L. (2023). *The ethics of artificial intelligence*. Oxford University Press.

FLORIDI, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1).

FOUTO, A. C., Rodrigues Rocha, F., Seixas, M., Brigas, M. A., & Silva Santos, J. (2018). *Textos de história do direito português*. AAFDL.

HESPANHA, A. M. (2012). *Cultura jurídica europeia: Síntese de um milénio*. Almedina.

HILDEBRANDT, M. (2015). *Smart technologies and the end(s) of law*. Edward Elgar.

HOLMES, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education*. Center for Curriculum Redesign.

HOLMES, W., & Porayska-Pomsta, K. (Eds.) (2023). *The ethics of artificial intelligence in education*. Routledge.

KAHLE, P., Colutto, S., Hackl, G., & Mühlberger, G. (2017). Transkribus: A service platform for transcription. *ICDAR*, pp. 19-24.

KATZ, D. M., Bommarito, M. J., Gao, S., & Arredondo, P. (2024). GPT-4 passes the bar exam. *Philosophical Transactions of the Royal Society A*, 382(2270), 20230254.

LOEVINGER, L. (1949). Jurimetrics: The next step forward. *Minnesota Law Review*, 33(5), pp. 455-493.

LUCKIN, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed*. Pearson.

MIAO, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO.

MOLLICK, E. (2024). *Co-intelligence: Living and working with AI*. Portfolio/Penguin.

PASQUALE, F. (2015). *The black box society*. Harvard University Press.

PASQUALE, F. (2020). *New laws of robotics*. Belknap Press.

READ-COOP. *Transkribus*. https://www.transkribus.org

RIEDLBERGER, P., & Rosenbaum, G. *Amanuensis*. https://www.riedlberger.de/amanuensis

RIEDLBERGER, P., & Rosenbaum, G. (2014 e seguintes). *Amanuensis* [software, base ROMTEXT, Universidade de Linz].

RUSSELL, S. (2019). *Human compatible*. Viking.

SAMPAIO E MELLO, G. de (1997). *Apontamentos para a história do ensino do direito romano em Portugal*. AAFDL.

SARTOR, G. (2020). *L'intelligenza artificiale e il diritto*. Giappichelli.

SEIXAS, M. (2016). *Pessoa e trabalho: Contributo para a história do direito do trabalho em Portugal (1750-1878)*. AAFDL.

SELWYN, N. (2019). *Should robots replace teachers?* Polity.

SUSSKIND, R. (2023). *Tomorrow's lawyers* (3.ª ed.). Oxford University Press.

TURING, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), pp. 433-460.

UNIVERSIDADE DE LISBOA. *Serviços digitais ULisboa* e *Serviços de inteligência artificial*. https://digital.ulisboa.pt e https://digital.ulisboa.pt/info/servicos-de-inteligencia-artificial

VASWANI, A., et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.

VERA-CRUZ PINTO, E. (2015). *O futuro da justiça*. Nova Vega.

VERA-CRUZ PINTO, E. (2020). *Curso de direito romano* (2.ª ed.). Principia.

VERA-CRUZ PINTO, E. (2022). *Introdução ao direito romano: As questões fundamentais*. AAFDL.

VIDEIRA, S. Antas (Coord.) (2009). *O perfil do juiz na tradição ocidental*. Almedina.

Recursos digitais e plataformas

IUS LUSITANIAE: Fontes históricas do direito português (Universidade Nova de Lisboa). http://iuslusitaniae.fcsh.unl.pt

NOTA SOBRE A UTILIZAÇÃO DE INTELIGÊNCIA ARTIFICIAL

Em coerência com a prática pedagógica que aqui se propõe, declara-se que este artigo recorreu a ferramentas de Inteligência Artificial em momentos da sua preparação, designadamente na revisão linguística e gramatical do texto e na organização estética do documento. Todas as fontes legais, doutrinárias e institucionais foram verificadas pelo autor nas respetivas fontes oficiais, a tese e os argumentos são da sua autoria e responsabilidade, e nenhuma passagem foi gerada por IA e apresentada como redação própria sem verificação. A declaração cumpre, ela mesma, o primeiro dos princípios defendidos no estudo: a transparência.

PERFIL CIENTÍFICO DO AUTOR

Para consulta do currículo completo e da produção científica do autor, acede-se ao seu perfil CIÊNCIAVITAE através do código seguinte.

https://www.cienciavitae.pt/951F-767D-3DF2



A REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL EM PORTUGAL: CONTRIBUTO DE *LEGE FERENDA* PARA UMA LEI DE EXECUÇÃO DO REGULAMENTO (UE) 2024/1689 (*EU AI ACT*)

FERNANDA DUARTE MARQUES*

* *PhDc* em Direito (Faculdade de Direito da Universidade de Lisboa e Universidade da Extremadura). MSc em Gestão (Instituto Superior de Gestão). Em exercício de funções na Polícia Judiciária, membro do Comité de Ética e Supervisão da IA da Polícia Judiciária e, na sua representação, especialista no *AI Experts Group* (*Law Enforcement and Security*) da Comissão Europeia. Professora convidada no Instituto da Polícia Judiciária e das Ciências Criminais. As opiniões expressas neste artigo são da exclusiva responsabilidade da autora, não refletindo, necessariamente, as eventuais posições das ligações institucionais detidas. Contacto: fduartemarquesmile@gmail.com.

RESUMO:

A Inteligência Artificial (IA) não é apenas uma tecnologia universal: ela está a reconfigurar toda a nossa economia, a nossa sociedade e o funcionamento das instituições. Logo, os desafios jurídicos que suscita exigem quadros normativos que garantam o seu desenvolvimento e utilização de modo seguro, fiável e ético, no respeito pelos nossos direitos fundamentais. O Regulamento (UE) 2024/1689 (RIA ou Regulamento), alterado pelo Regulamento (UE) 2026/1744 (Omnibus Digital sobre IA), estabeleceu um regime harmonizado baseado no risco, mas carece de medidas nacionais de execução para a sua aplicação efetiva elencadas no próprio Regulamento. Em diversos Estados-Membros, o respetivo processo legislativo já se encontra em curso ou está mesmo concluído, pelo que urge iniciá-lo em Portugal. Daí o contributo *de lege ferenda* ínsito no presente artigo, acompanhado da motivação das opções perfilhadas na proposta apresentada de anteprojeto de lei de execução do Regulamento. No âmbito da governação, consagra-se a designação da ANACOM como ponto de contacto único e autoridade de fiscalização do mercado, apoiada por um Conselho de Coordenação e um Registo Nacional de Sistemas de IA. Ponderando as exigências de liberdade e de segurança, densificam-se os pressupostos e as garantias da utilização, para efeitos de aplicação da lei, de sistemas de identificação biométrica à distância, em tempo real e em diferido. Contempla-se, ainda, para estimular a inovação tecnológica, um ambiente de testagem da regulamentação. Paralelamente, a violação das normas e o incumprimento das obrigações preconizadas no Regulamento e na futura lei de execução impõem que seja definido um regime sancionatório articulado com a proteção de dados, os serviços digitais e a supervisão financeira, respeitando o princípio ne bis in idem. Pretende-se, assim, com este contributo, que a futura lei, concebida à luz das especificidades do ordenamento jurídico português, equilibre, jurídica e eticamente, a promoção da inovação e a salvaguarda efetiva dos direitos fundamentais.

Palavras-chave: Regulamento da inteligência artificial; Anteprojeto de lei de execução; Governação da IA; identificação biométrica; ambientes de testagem; regime sancionatório.

ABSTRACT:

Artificial intelligence (AI) is not merely a universal technology: it is reconfiguring our entire economy, society, and the functioning of institutions. Therefore, the legal challenges it raises require regulatory frameworks that ensure its development and use in a safe, reliable, and ethical manner, with respect for our fundamental rights. Regulation (EU) 2024/1689 (RIA or Regulation), as amended by Regulation (EU) 2026/1744 (Digital Omnibus on AI), established a harmonised risk-based regime, but lacks national execution measures for its effective application set out in the Regulation itself. In several Member States, the respective legislative process is already underway or even concluded, which is why it is urgent to initiate it in Portugal. Hence the *de lege ferenda* contribution embedded in this article, accompanied by the rationale for the choices made in the presented draft bill executing the Regulation. Within the scope of governance, the designation of ANACOM as the single point of contact and market surveillance authority is established, supported by a Coordination Council and a National Register of AI Systems. Balancing the demands of freedom and security, the conditions and safeguards for the use of remote biometric identification systems, both real-time and post-remote, for law enforcement purposes are detailed. To stimulate technological innovation, a regulatory sandbox is also envisaged. In parallel, the violation of rules and non-compliance with the obligations set forth in the Regulation and the future executing law require the establishment of a sanctioning regime coordinated with data protection, digital services, and financial supervision, while respecting the ne bis in idem principle. The objective of this contribution is thus for the future law, designed considering the specificities of the Portuguese legal system, to achieve a legal and ethical balance between promoting innovation and effectively safeguarding fundamental rights.

Keywords: Artificial Intelligence Act; Draft implementing law; AI governance; biometric identification; sanctioning regime; regulatory sandboxes.

III. NÓTULA INTRODUTÓRIA

A Inteligência Artificial (IA) não é uma realidade emergente, ainda que a sua popularidade, transversalidade e globalização o sejam, em especial nos últimos anos. Na verdade, a IA assenta num longo percurso histórico: adquiriu reconhecimento e autonomia científica há já sete décadas, na célebre Conferência de Dartmouth de 1956, na qual foi consagrada, pela primeira vez, a locução *"inteligência artificial"*. As décadas seguintes foram de desenvolvimento assíncrono, com etapas de acentuado progresso, intercaladas por períodos de desinvestimento e estagnação, os designados *"invernos da IA"*. Com a viragem do século, a disponibilidade de grandes volumes de dados (*big data*), o aumento da capacidade computacional e o desenvolvimento da aprendizagem profunda (*deep learning*) impulsionaram decisivamente a evolução desta tecnologia. A partir de 2020, o aparecimento dos grandes modelos de linguagem (*Large Language Models – LLM*) e da IA generativa, em particular o lançamento do ChatGPT em 2022, marca a abertura de uma nova fase[1]. Esta carateriza-se pela existência de sistemas capazes de gerar conteúdos, interagir com os utilizadores e executar tarefas progressivamente mais complexas e autónomas, bem como pela democratização de uma tecnologia hoje acessível a todos. Daí a atenção atribuída à IA pelos governos nacionais e pelas organizações internacionais, que ponderam a evolução da inovação tecnológica, com inegáveis ganhos de eficiência, com a proteção dos direitos fundamentais[2] face aos riscos associados[3]. O intuito é que a IA seja utilizada de modo seguro,

1 Sobre a evolução histórica da IA, *vd*, em especial, NILSSON, Nils J., *The quest for artificial intelligence: a history of ideas and achievements.* Cambridge: Cambridge University Press, 2010, consultado em 25 de maio de 2026. Disponível em: https://ai.stanford.edu/~nilsson/QAI/qai.pdf e RUSSELL, Stuart, J. e NORVIG, Peter, *Artificial Intelligence: a modern approach*. Pearson: Pearson Education, 2021, pp. 5-30.

2 Sobre este tema, *vd.* CASTRO, Raquel Brízida, *Direito Constitucional - Ciberespaço e Tecnologia*. Coimbra: Almedina, 2023; MONIZ, Graça Canto e VISITAÇÃO, Inês, Um primeiro olhar sobre a avaliação de impacto no regulamento geral sobre a proteção de dados (RGPD) e no regulamento da inteligência artificial (RIA*), AI Act - o novo regulamento de inteligência artificial europeu.* Coimbra: Almedina, 2025, pp.457-472; KARATHANASIS, Theodoros, The FRIA in EU AI Act: governance, rights and global jurisdiction, Law, Innovation and Technology, 2026, pp. 1-24. Consultado em 25 de maio de 2026. Disponível em: https://www.tandfonline.com/doi/full/10.1080/17579961.2026.2710952; FARINHO, Domingos Soares, Os Direitos Fundamentais no Regulamento da Inteligência Artificial: levando os direitos fundamentais a sério ou manobra de diversão?*, AI Act - o novo regulamento de inteligência artificial europeu.* Coimbra: Almedina, 2025, pp.545-567.

3 No âmbito desses riscos, destacam-se o enviesamento algorítmico, a opacidade (*black-box*), a discriminação, a manipulação, a perda de privacidade e de controlo humano. *Vd.*, entre outros, VEALE, Michael e BORGESIUS, Frederik, Demystifying the Draft EU Artificial Intelligence Act: analysing the good, the bad, and the unclear elements of the proposed approach, *Computer Law Review International*, 4/2021. Consultado em 15 de maio de 2026. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3896852;

fiável e ético, em benefício da economia, das instituições, da sociedade e, sobretudo, dos cidadãos.

É neste contexto que a IA é inscrita na agenda europeia, e num momento de intensa concorrência internacional com os EUA e a China. A estratégia[4] definida, acompanhada de instrumentos de *soft law*[5], conduziu à apresentação, em 2021, de uma proposta de regulamento[6]. Após um amplo debate e um processo legislativo demorado, essa proposta resultou no Regulamento (EU) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024 (Regulamento Europeu da IA, o designado *RIA*)[7]. O seu regime foi, posteriormente, alterado pelo Regulamento (EU) 2026/1744, de 8 de julho de 2026 (*Omnibus*

RAJENDRA, Josephine Bhavani e THURAISINGAM, Ambikai S., The role of explainability and human intervention in AI decisions: jurisdictional and regulatory aspects, *Information & Communications Technology Law*, 35:2, 152-183. Consultado em 15 de maio de 2026. Disponível em: https://www.tandfonline.com/doi/full/10.1080/13600834.2025.2537514?scroll=top&needAccess=true; BARROSO, Luís Roberto e MELLO, Patrícia Perrone Campos, "*Artificial Intelligence: Promises, risks, and regulation. Something new under the sun*", *Harvard Latin American Law Review*, 2025, vol. 28, pp. 73-102.

4 COMISSÃO EUROPEIA, Comunicação da Comissão ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões: *Inteligência Artificial para a Europa.* COM(2018) 237 final. Bruxelas: Comissão Europeia, 2018. Consultado em 25 de maio de 2026. Disponível em: https://eur-lex.europa.eu/ ; COMISSÃO EUROPEIA, Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões: *Edificar a Confiança numa Inteligência Artificial Centrada no Ser Humano*. COM(2019) 168 final. Bruxelas: Comissão Europeia, 2019. Consultado em 25 de maio de 2026. Disponível em: https://digital-strategy.ec.europa.eu/en/library/communication-building-trust-human-centric-artificial-intelligence

5 Entre outros, GRUPO DE PERITOS DE ALTO NÍVEL SOBRE A IA, Orientações éticas para uma IA de confiança. Luxemburgo: Serviço de Publicações da UE, 2019. Consultado em 25 de maio de 2026. Disponível em: https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-pt.

6 COMISSÃO EUROPEIA, *Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial*, de 21 de abril de 2021 (Regulamento Europeu da Inteligência Artificial). COM(2021) 206 final. Bruxelas: Comissão Europeia, 2021. Consultado em 25 de maio de 2026. Disponível em: https://data.consilium.europa.eu/doc/document/ST-8115-2021-INIT/pt/pdf

7 O *EU AI Act* foi amplamente criticado, pelos mais diversos setores e personalidades, que argumentaram que a complexidade, fragmentação e sobreposição regulatórias, a onerosidade e subsequente morosidade decorrente do *compliance* dos sistemas de IA de risco elevado, os prazos exíguos acompanhados de um severo quadro sancionatório eram fatores de perda de competitividade europeia quanto à IA. *Vd.* DRAGHI, Mario, *The Future of European Competitiveness: Report by Mario Draghi*. Bruxelas: Comissão Europeia, 2024. Consultado em 6 de julho de 2026. Disponível em: https://commission.europa.eu/topics/strengthening-european-competitiveness/draghi-report_en; KILIAN R. et al., European AI Standards – Technical Standardisation and Implementation Challenges under the EU AI Act*, European Journal of Risk Regulation*. 2025; 16(3); pp. 1038-1062. Consultado em 6 de julho de 2026. Disponível em: https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/european-ai-standards-technical-standardisation-and-implementation-challenges-under-the-eu-ai-act/E5157BA0391FFA9E1A3233E636005192; SCHREPEL, Thibault, Decoding the AI Act: Implications for Competition Law and Market Dynamics. *Journal of Competition Law & Economics*, Oxford, v. 21, n. 3, p. 374-389, 2025. Consultado em 6 de julho de 2026. Disponível em: https://academic.oup.com/jcle/article/21/3/374/8051960.

Digital sobre IA)[8], de modo a introduzir medidas de simplificação, ajustar o calendário de aplicação e consagrar novas práticas proibidas de IA. Embora o Regulamento Europeu da IA, por força da sua natureza jurídica, tenha aplicabilidade direta em todos os Estados-Membros, nos termos do artigo 288.º do Tratado sobre o Funcionamento da União Europeia, não dispensa a intervenção do legislador nacional, expressamente prevista no próprio Regulamento. Cabe, nomeadamente, aos Estados-Membros estruturar a governação da IA, criar pelo menos um ambiente de testagem da regulamentação, disciplinar a utilização de sistemas de identificação biométrica à distância para efeitos de aplicação da lei e estabelecer o regime sancionatório. Vários Estados-Membros já iniciaram esse processo legislativo, a ritmos diferentes e com opções nacionais distintas. Portugal terá igualmente de o fazer.

O presente artigo propõe-se, assim, contribuir para a futura lei de execução do Regulamento Europeu da Inteligência Artificial[9], de modo que esta, concebida à luz das especificidades do ordenamento jurídico português, equilibre, jurídica e eticamente, a promoção da inovação e a salvaguarda efetiva dos direitos fundamentais. Ademais, a futura lei de execução nacional desempenhará um papel decisivo na gestão de um complexo pluralismo normativo, de perfil multinível, no qual confluem um vasto acervo de disciplinas jurídicas, de ramos distintos,

8 O *Omnibus Digital* sobre IA não foi isento de críticas, quer por reconfigurar a governança digital europeia a favor da competitividade e da aceleração da IA, com prejuízo das salvaguardas democráticas e dos direitos fundamentais, *vd.* FLORES MOLEÓN, Alejandro e OLEART, Alvaro. "*When AI-First Becomes Democracy-Last: The European Commission's Digital Omnibus and Its Technosolutionism on Steroids*". *European Journal of Risk Regulation*, Cambridge, First View, p. 1-15, 2026. Consultado em 6 de julho de 2026. Disponível em: https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/when-aifirst-becomes-democracylast-the-european-commissions-digital-omnibus-and-its-technosolutionism-on-steroids/06977360816348D2DB915A8FDDD34CEB; quer por aumentar riscos para os direitos fundamentais e para a segurança jurídica, sem efetivamente simplificar o quadro regulatório, vd. RUSCHEMEIER, Hannah; DE GREGORIO, Giovanni. "*From a Right-Based Approach to Competitiveness? The EU Digital Omnibus Reform*". *European Journal of Risk Regulation*, Cambridge, First View, p. 1-18, 2026. Consultado em 6 de julho de 2026. Disponível em: https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/from-a-rightbased-approach-to-competitiveness-the-eu-digital-omnibus-reform/FBC0D59ECA30BD6AE32C5622367BB4D6

9 O *EU AI Act*, sendo um regulamento europeu, é de aplicabilidade direta nos Estados-Membros, mas a sua efetividade depende de medidas de execução, em algumas matérias, que o próprio Regulamento consigna (governação, identificação biométrica à distância para efeitos da aplicação da lei, regime sancionatório). Sobre esta articulação e os desafios que urgem da mesma, *vd.* PAOLUCCI, Federica, "*The enforcement of the Artificial Intelligence Act: Looking Forward to a Commission Implementing Decisions for Protecting Fundamental Rights*". *SSRN*. Consultado em 6 de julho de 2026. Disponível em: https://ssrn.com/abstract=5147813; NEVES, Inês, "*Apontamentos de Governance e de Enforcement no Regulamento de Inteligência Artificial: sobre a dimensão nacional de um modelo misto. AI Act - o novo regulamento de inteligência artificial europeu*". Coimbra: Almedina, 2025, pp. 125-173.

bem como de autoridades, quer na esfera europeia, quer na ordem jurídica interna[10]. A intervenção do legislador nacional permite, deste modo, regular a articulação entre esses diversos regimes jurídicos e atribuições institucionais, assegurando a respetiva delimitação e a sua harmonia com a proteção dos direitos em apreço.

IV. REGULAMENTO EUROPEU DA INTELIGÊNCIA ARTIFICIAL: TRAÇOS ESSENCIAIS[11]

O Regulamento Europeu da Inteligência Artificial constitui o primeiro quadro normativo horizontal dedicado aos sistemas de IA, abrangendo o seu desenvolvimento, colocação no mercado e respetiva utilização. O legislador procurou equilibrar o funcionamento do mercado interno com a proteção da saúde, da segurança e dos direitos fundamentais, promovendo a inovação tecnológica, assentando na premissa de uma IA confiável e centrada no ser humano, conforme expressa no artigo 1.º do RIA. Daqui resulta um regime híbrido, que combina mecanismos da segurança dos produtos, como a avaliação da conformidade ou a fiscalização do mercado, com exigências de tutela dos direitos fundamentais. Esta opção foi objeto de crítica pela doutrina[12], que alerta que esta combinação, se ignorar as diferenças estruturais entre as duas tradições, pode conduzir a uma proteção meramente formal dos direitos envolvidos. Em seguida, procede-se a uma análise sucinta dos traços essenciais do RIA.

---

10 *Vd.* CASTRO, Raquel Brízida, op. cit., pp. 84, esta autora designa este pluralismo normativo, concernente à esfera europeia, como *"prolixidade normativa"*. Destaca que a *"monumental dispersão jurídica, normativa e institucional, propiciadora de uma perturbante fragmentação normativa"* aliada ao ritmo das inovações tecnológicas é potenciadora de *"gerar insegurança jurídica em áreas constitucionalmente protegidas"*, a que acresce *"a previsão a criação ou alocação de vários Reguladores, europeus e nacionais, cada um com o seu perfil, natureza jurídica, modelo institucional e grau de independência"*.

11 Sobre o EU AI Act, nos seus diversos domínios de aplicação, vd. RODRIGUES, André Alfar et al, (coord). "*AI Act - O novo regulamento de inteligência artificial europeu*". Coimbra: Almedina, 2025.

12 ALMADA, Marco, PETIT, Nicolas, "*The EU AI act: a medley of product safety and fundamental rights?*", EUI, RSC, Working Paper, 2023/59. Consultado a 15 de maio de 2026. Disponível em: https://hdl.handle.net/1814/75982; ALMADA, Marco, PETIT, Nicolas, "*The EU AI Act: Between the rock of product safety and the hard place of fundamental rights*", *Common Market Law Review*, Volume 62, Issue 1 (2025) pp. 85 - 120. Consultado em 15 de maio de 2026. Disponível em: https://kluwerlawonline.com/journalarticle/Common+Market+Law+Review/62.1/COLA2025004

a) Âmbito de aplicação

Relativamente ao seu âmbito de aplicação, nos termos do artigo 2.º, n.º 1, o RIA aplica-se aos prestadores, responsáveis pela implantação, importadores, distribuidores, fabricantes de produtos e mandatários, bem como às pessoas afetadas que se encontrem na EU. Quanto aos prestadores que coloquem sistemas de IA no mercado ou em serviço na EU, é indiferente o local de estabelecimento, que pode situar-se fora da EU; e os prestadores e responsáveis pela implantação estabelecidos em países terceiros ficam igualmente abrangidos quando os resultados do sistema sejam utilizados na EU. Releva, pois, a conexão com o mercado europeu, à semelhança do Regulamento Geral de Proteção de Dados (RGPD)[13]. Excluem-se, entre outros, os usos exclusivamente militares, de defesa ou de segurança nacional, a investigação científica e a utilização pessoal, mas permanecem abrangidas as autoridades policiais e judiciárias (n.ºs 3, 4, 6, 8, 10, 12).

b) Definição de sistema de IA

O RIA consagra, no seu artigo 3.º, um vasto elenco de sessenta e oito definições legais, destacando-se a qualificação de *"sistema de IA"*. Nos termos do seu ponto 1), este é preconizado como *"um sistema baseado em máquinas concebido para funcionar com níveis de autonomia variáveis, e que pode apresentar capacidade de adaptação após a implantação e que, para objetivos explícitos ou implícitos, e com base nos dados de entrada que recebe, infere[14] a forma de gerar resultados, tais como previsões, conteúdos, recomendações ou decisões que podem influenciar ambientes físicos ou virtuais"*[15]. A opção pelo estabelecimento de uma definição legal encerra um dilema normativo estrutural: se, por um lado, desempenha uma função de delimitação na sua aplicação, o que promove a segurança jurídica e uniformidade hermenêutica, desde que não haja nem demasiada amplitude nem

13 *Vd.* artigo 3.º do RGPD.

14 O *EU AI Act,* como elemento central da definição, consagra a inferência, ou seja, a capacidade de gerar resultados autonomamente, para distinguir a IA do software tradicional sem identificar tecnologias específicas, mas têm sido formuladas críticas, dada as ambiguidades jurídicas e técnicas no termo que dificultam a sua aplicação a novos modelos (*v.g.* se a inferência é no desenvolvimento ou na implementação ou a delimitação concreta entre software tradicional avançado e IA). *Vd.* MENA-DURÁNA, Maria L., e MONTESINOS, Antonio, "*Inference as the defining criterion of AI: rethinking the EU AI Act*", *Law, Innovation and Technology, 2026,* pp. 1-27. Consultado em 20 de maio de 2026. Disponível em: *https://www.tandfonline.com/doi/full/10.1080/17579961.2026.2705130*

15 Esta definição foi influenciada pela definição atual preconizada pela OCDE, na sua *Recommendation of the Council on Artificial Intelligence*. Consultada em 20 de maio de 2026. Disponível em: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

estreiteza; por outro lado, potencia os riscos de rigidez e obsolescência regulatória face ao perfil dinâmico do direito digital, e ainda mais da IA[16]. Esta definição tem sido objeto de vincada crítica doutrinária que a considera excessivamente ampla e indeterminada[17], com prejuízo para a segurança jurídica, dificuldade que as orientações emitidas pela Comissão Europeia atenuam, mas não resolvem em pleno[18].

c) Abordagem baseada no risco

O RIA gradua as obrigações em função do risco, com quatro níveis: risco inaceitável (práticas proibidas, as quais estão previstas no artigo 5.º), risco elevado (artigo 6.º e seguintes), risco limitado (sujeitos a deveres de transparência, nos termos do artigo 50.º) e risco mínimo ou nulo (sem sujeição a obrigações específicas, mas com o incentivo da adoção de códigos de conduta, em conformidade com o disposto no artigo 95.º), a que acresce um regime próprio para os modelos de finalidade geral (artigos 51.º e seguintes). Esta opção tem sido objeto de amplas críticas, nomeadamente quer pela (in)adequação do paradigma do risco à tutela de direitos individuais[19], quer pelo nteligê estático de uma classificação estabelecida *ex ante*[20].

Relativamente às práticas proibidas, estas correspondem a utilizações de sistemas de IA que o legislador europeu considerou incompatíveis com os valores da EU, como sejam a

16 Sobre esta tensão dialética entre segurança jurídica e fixação do conceito e a necessidade da sua adaptabilidade à velocidade da evolução tecnológica (*"pacing problem"*), *vd.* LARSSON, Stefan, et al,, "*Implications of regulating a moving target: between fixity and flexibility in the EU AI Act*", Law, Innovation and Technology 18.1,, 83-106. Consultado em 20 de maio de 2026, disponível em: https://www.tandfonline.com/doi/full/10.1080/17579961.2026.2633682 Estes autores também alertam para o facto de a flexibilidade excessiva poder reduzir a previsibilidade legal e transferir um poder (demasiado) amplo para comissões técnicas e entidades de certificação.

17 MEINEL, Philip et al., "*What are AI systems? Rethinking the core definition in the EU AI Act*". In: *Computer Law & Security Review*, Vol. 62, No. 106373, Pp. 1-12, Elsevier, 2026. Consultado em 20 de maio de 2026. Disponível em: https://www.dfki.de/fileadmin/user_upload/import/17530_What_are_AI_systems__Rethinking_the_core_definition_in_the_EU_AI_Act.pdf

18 COMISSÃO EUROPEIA, 2025. Commission *Guidelines on the definition of an AI System established by Regulation (EU) 2024/1689*. C(2025) 5053 final. Bruxelas: Comissão Europeia, 2025. Consultado em 20 de maio de 2026. Disponível em: https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application

19 KAMINSKI, Margot E., "*Regulating the Risks of AI*" (August 19, 2022). *Boston University Law Review*, Vol. 103:1347, 2023, U of Colorado Law Legal Studies Research Paper No. 22-21. Consultado em 15 de maio de 2026. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4195066

20 NOVELLI, C. *et al.* "*Taking AI risks seriously: a new assessment model for the AI Act*". *AI & Soc* 39, 2493–2497 (2024). Consultado em 15 de maio de 2026. Disponível em: https://link.springer.com/article/10.1007/s00146-023-01723-z

manipulação e exploração de vulnerabilidades causadoras de danos significativos, a classificação social, a avaliação do risco criminal baseada apenas em perfis, a recolha aleatória de imagens faciais, o reconhecimento de emoções no trabalho e no ensino, a categorização biométrica de caraterísticas sensíveis e, salvo exceções taxativas e com autorização prévia, a identificação biométrica à distância em tempo real para efeitos da aplicação da lei. O *Omnibus Digital* sobre IA aditou a proibição de sistemas que gerem imagens íntimas não consentidas ou material de abuso sexual de crianças[21]. A Comissão Europeia publicou orientações interpretativas sobre este artigo[22], bastante desenvolvidas e ilustrativas.

No que concerne aos sistemas de risco elevado, é aqui que reside o cerne do regime. Nos termos do artigo 6.º, são de risco elevado os sistemas de IA integrados em produtos regulados pelo Anexo I e os utilizados nos domínios do Anexo III, entre os quais a biometria, o emprego, o acesso a serviços essenciais, a aplicação da lei, a migração e a justiça, salvo quando, nos termos do n.º 3, não representem risco significativo. A amplitude destes domínios e das obrigações associadas explanam a sua relevância. Os prestadores suportam a maior parte das obrigações, desde a gestão de riscos à supervisão humana e à avaliação da conformidade (artigos 9.º a 17.º). Os mandatários, importadores e distribuidores têm deveres de verificação e cooperação (artigos 22.º a 24.º) e os responsáveis pela implantação devem assegurar a utilização conforme (artigo 26.º) e, nos casos previstos no artigo 27.º, avaliar o impacto sobre os direitos fundamentais. A Comissão Europeia publicou, em maio de 2026, um projeto de orientações sobre os sistemas de risco elevado[23], à semelhança do que tinha feito nomeadamente para a definição do sistema de IA e das práticas proibidas[24].

---

21 A Comissão Europeia responde, assim, a uma necessidade urgente de obviar (ou, pelo menos mitigar) os riscos críticos e em expansão e crescentes associados à utilização de ferramentas de IA generativa (*vd*. Considerandos 10 a 16 do *Omnibus Digital* sobre IA).

22 COMISSÃO EUROPEIA, Commission *Guidelines on prohibited AI practices established by Regulation (EU) 2024/1689 (AI Act)*. C(2025) 5052 final. Bruxelas: Comissão Europeia, 2025. Consultado em 15 de maio de 2026. Disponível em: guidelines_on_prohibited_artificial_intelligence_practices_established_by_regulation_eu_20241689_ai_act_english_ied3r5nwo50xggpcfmwckm3nuc_112367-1.PDF

23 COMISSÃO EUROPEIA, 2026. *Draft Commission Guidelines on the Classification of High-Risk AI Systems under Article 6 of Regulation (EU) 2024/1689 (AI Act).* Bruxelas: Comissão Europeia, 19 maio 2026. Consultado em 15 de maio de 2026. Disponível em: https://digital-strategy.ec.europa.eu/en/library/draft-commission-guidelines-classification-high-risk-ai-systems

24 Embora estes já em versão final.

d) Calendário de aplicação

O RIA entrou em vigor a 1 de agosto de 2024, sendo aplicado de modo faseado (artigo 113.º): as práticas proibidas e a literacia desde 2 de fevereiro de 2025, os modelos de finalidade geral e as sanções desde 2 de agosto de 2025, a generalidade das restantes normas desde 2 de agosto de 2026. O *Omnibus Digital* sobre IA, em vigor desde 27 de julho de 2026, adiou as obrigações de risco elevado para 2 de dezembro de 2027 (Anexo III) e 2 de agosto de 2028 (Anexo I).

O legislador europeu optou por uma implementação faseada do Regulamento, na nossa opinião plenamente justificada, de modo a assegurar *um "período de adaptação adequado para os operadores e evitar perturbações do mercado"*[25], e, mesmo assim, teve de ser reajustada com o *Omnibus Digital* sobre IA.

e) Ambientes de testagem de regulamentação

A Comissão Europeia recorreu a um instrumento usado no setor financeiro (*regulatory sandbox*) e adaptou-o para a inovação tecnológica. As *regulatory sandboxes* são "*regimes criados por autoridades competentes que permitem a entidades reguladas e não reguladas testar, nos termos de um plano de testes previamente acordado e monitorizado por um setor dedicado dessas autoridades competentes – produtos, serviços, modelos de negócio e mecanismos de distribuição que sejam inovadores no setor relevante*"[26]. É um mecanismo para promover a inovação tecnológica, em que cada Estado-Membro está obrigado a criar pelo menos um ambiente de testagem, onde sistemas inovadores podem ser desenvolvidos e testados sob supervisão, nos termos dos artigos 57.º a 63.º do RIA. O prazo de criação foi prorrogado para 2 de agosto de 2027 pelo *Omnibus Digital* sobre IA, o que permitirá aos Estados-Membros usarem estes ambientes de testagem de regulamentação (*regulatory sandboxes*) com base na experiência existente no setor financeiro.

25 *Vd.* Considerando 177.

26 CRAVO, Laura Abreu e SANTOS, João Vieira, "*Regulatory sandboxes. Cadernos do Mercado de Valores Mobiliários sobre os 20 anos da CMVM*". Coimbra: Almedina, pp. 639-656, 641. *Vd.,* também, PLATO-SHINAR, Ruth e GODWIN, Andrew, "*Regulatory cooperation in AI sandboxes: insights from Fintech*". Law, Innovation and Technology, vol. 18, nr. 1, pp. 320-347. Consultado em 7 de julho de 2026. Disponível em: https://www.tandfonline.com/doi/full/10.1080/17579961.2026.2633675.

f) Regime sancionatório

Para garantir a eficácia do Regulamento, o artigo 99.º do RIA estabelece os limites máximos das coimas aplicáveis às infrações, cabendo aos Estados-Membros estabelecer o respetivo regime sancionatório, incluindo a decisão sobre se, e em que medida, podem ser aplicadas coimas a autoridades e organismos públicos[27]. Esses limites são graduados em função da gravidade da infração: até 35 milhões de euros ou 7% do volume de negócios anual a nível mundial, consoante o montante mais elevado, pela violação das práticas proibidas previstas no artigo 5.º; até 15 milhões de euros ou 3% pelo incumprimento das demais obrigações dos operadores e dos organismos notificados; e até 7,5 milhões de euros ou 1% pela prestação de informações incorretas, incompletas ou enganosas aos organismos notificados ou às autoridades nacionais competentes.

Conforme decorre dos considerandos 168 e seguintes, o cumprimento das normas exige sanções efetivas, proporcionadas e dissuasivas, cuja determinação deve atender à natureza, gravidade e duração da infração, bem como às suas consequências, designadamente para os direitos fundamentais. Por sua vez, para salvaguardar a viabilidade das PMEs, incluindo as empresas em fase de arranque, o legislador determinou que lhes seja aplicado o menor dos montantes ou percentagens previstas.

O exercício do poder sancionatório fica, em qualquer caso, sujeito a garantias procedimentais e processuais adequadas e ao respeito pelo princípio *ne bis in idem*[28].

De notar que os limites anteditos superam os do RGPD (20 milhões de euros ou 4%, conforme previsto no artigo 83.º, n.º 5), refletindo a perceção de que os riscos da IA para os direitos fundamentais, em especial as práticas proibidas, justificam um efeito dissuasivo mais severo.

27 O que poderá ser potenciador de uma dupla assimetria: entre operadores públicos e privados no mesmo Estado e, entre Estados-Membros, consoante a opção de cada um.

28 Sobre esta matéria, *vd.* SAAVEDRA, Alberto, As coimas do Regulamento de Inteligência Artificial: A perspetiva nacional do enforcement. *AI Act – O novo Regulamento de Inteligência Artificial Europeu.* Coimbra: Almedina, 2025, pp. 475-494.

V. EXPERIÊNCIAS COMPARADAS

Conforme foi expresso anteriormente, o Regulamento Europeu da IA, por força da sua natureza jurídica, tem aplicabilidade direta em todos os Estados-Membros, mas a sua execução, para efeitos de uma aplicação efetiva, necessita de medidas legislativas nacionais, de modo a estruturar a governação da IA, a criar um ambiente de testagem da regulamentação, a definir a utilização de sistemas de identificação biométrica à distância para efeitos de aplicação da lei e a estabelecer o regime sancionatório.

A opção europeia adquire relevância quando comparada com as dos outros grandes polos de desenvolvimento da IA: EUA, Japão e China, cuja convergência global é assumida pela doutrina como inviável[29]. A EU é a única com um ato legislativo horizontal vinculativo e diretamente aplicável. O *RIA* inclui práticas proibidas, obrigações consoante os níveis de risco, avaliação de conformidade e sanções. A China perfilhou um modelo vertical, impondo regras obrigatórias, mas dispersas por diplomas setoriais, versando, entre outros, sobre IA generativa e revisão ética. Tanto na EU como na China, as sanções são fixadas administrativamente. Nos EUA, a regulação federal depende de atos executivos sucessivamente revogados e de estatutos setoriais anteriores à IA. O Japão afastou expressamente requisitos horizontais vinculativos. Em termos de risco, apenas a EU consagra uma tipologia legal de risco em quatro níveis. Os EUA conhecem uma diferenciação semelhante, sem força jurídica. Na China, a classificação é implícita e assenta em juízos sobre harmonia social e segurança. O Japão confere o poder de avaliação às empresas. De acordo com Cooper e Kompella[30], a regulação de IA é uma afirmação geopolítica, em que cada pólo legisla em função de um interesse estratégico próprio. A EU procura a soberania digital, reduzindo a dependência tecnológica de terceiros, os EUA pretendem assegurar a liderança tecnológica face à China, esta utiliza a regulação para assegurar que os sistemas de IA estão em harmonia com a ideologia do Estado e, por fim, o Japão enfatiza um quadro flexível para atrair investimento e talento. Resultam, assim, dois modelos, prosseguem os

29 *Vd.* COOPER, J. M. e KOMPELLA, K, "*A Comparative View of Artificial Intelligence Regulation in the European Union*", Japan, People's Republic of China, and the United States of America. *Arizona Journal of International and Comparative Law*, 2026, 43(1), 42-71. Consultado a 7 de julho de 2026. Disponível em: https://scholarlycommons.law.cwsl.edu/fs/538/

30 COOPER, J. M. e KOMPELLA, K,, *op. cit.*, pp.67-69.

autores. A EU e a China seguem uma abordagem dirigista, em que o Estado intervém ativamente e fixa regras vinculativas, mas que assentam em fundamentos diversos. Na EU assenta na proteção dos direitos fundamentais e na China no controlo estatal. Os EUA e o Japão privilegiam a competitividade, mediante a autorregulação, complementada por orientações não vinculativas e intervenções públicas pontuais. Ambas as abordagens comportam limitações estruturais, enquanto o modelo dirigista pode incorrer em sobrerregulação, o que pode desincentivar a inovação, o modelo orientado pelo mercado tende para a sub-regulação, comprometendo a salvaguarda contra riscos e danos sociais.

Por sua vez, na EU, vários Estados-Membros já iniciaram o processo legislativo de execução do *RIA*. Para efeitos da presente análise, foram selecionados dois ordenamentos jurídicos em fases distintas. Itália dispõe de um regime já concluído, a *Legge 23 settembre 2025, n. 132*, em vigor desde 10 de outubro de 2025. Em Espanha, o processo legislativo encontra-se ainda em curso, com o *Proyecto de Ley Orgánica para el buen uso y la gobernanza de la nteligência artificial* em tramitação parlamentar desde junho de 2026. A comparação permite observar tanto uma solução consolidada como as opções que estão a ser debatidas, e extrair daí elementos úteis para o legislador português.

a) Itália

Em Itália foi aprovada a *Legge 23 settembre 2025, n. 132*[31], cujo artigo 1.º define como objeto o estabelecimento de princípios em matéria de investigação, experimentação, desenvolvimento, adoção e aplicação de sistemas e modelos de IA, a promoção de uma utilização correta, transparente e responsável, numa dimensão antropocêntrica, e a fiscalização dos riscos económicos e sociais e do impacto de IA nos direitos fundamentais. Não se trata de uma lei meramente de execução do *RIA*, enuncia princípios gerais (artigo 3.º) e disciplina domínios setoriais, entre os quais a saúde, o trabalho, a administração pública e a justiça.

No âmbito da governação, a lei designa como autoridades nacionais a *Agenzia per l'Italia digitale* (AgID), autoridade notificadora, e a *Agenzia per la cybersicurezza nazionale* (ACN), como autoridade de fiscalização do mercado e ponto de contacto único. As autoridades

31 ITÁLIA, Legge 23 settembre 2025, n.º 132, *Gazzetta Ufficiale delaa Repubblica Italiana*. Serie generale, n.º 223. Consultado a 8 de julho de 2026. Disponível em: https://www.gazzettaufficiale.it/eli/id/2025/09/25/25G00143/sgMASE

setoriais mantêm as suas atribuições, e junto da Presidência do Conselho de Ministros funciona um Comité de Coordenação (artigos 19.º e 20.º). Esta lei delegou no Governo, pelo artigo 24.º, a adaptação ao *RIA* e a disciplina jurídica dos usos ilícitos de IA, delegação concretizada em dois diplomas.

O *Decreto Legislativo 7 ottobre 2026, n. 179*[32], em vigor a partir de 23 de outubro de 2026, desenvolve a governação e o regime sancionatório. As coimas, previstas no artigo 23.º, seguem os limites do artigo 99.º do RIA. Os artigos 26.º a 31.º regulam o *spazio di sperimentazione italiano per l'IA*, ambiente de testagem de regulamentação de IA, gerido pela AgID e pela ACN e com prioridade para PMEs e empresas em fase de arranque *(start-ups)*.

O *Decreto Legislativo 9 settembre 2026*, n. 160[33], em vigor desde 30 de setembro de 2026, regula o uso da IA pelas forças policiais. Para fins de investigação, a identificação biométrica à distância em tempo real depende de autorização do juiz de instrução, por 15 dias prorrogáveis, nos termos do novo artigo 359-ter do *nteli di procedura penale*, introduzido pelo artigo 13.º. Em situação de urgência, o *pubblico ministero* comunica o despacho ao juiz em 24 horas, e este valida-o em 48 horas. O mesmo artigo proíbe a utilização de bases de dados biométricas constituídas através da recolha aleatória de imagens faciais. O artigo 9.º exige avaliações de impacto sobre os direitos fundamentais e sobre a proteção de dados, enquanto o artigo 10.º sujeita o reconhecimento facial diferido, para pesquisa dirigida de suspeitos, a autorização judicial.

b) Espanha

Logo em 2023, antes da entrada em vigor do *RIA*, o *Real Decreto* 729/2023[34], de 22 de agosto, criou a *Agencia Española de Supervisión de Inteligencia Artificial* (AESIA), com

32 ITÁLIA, Decreto legislativo 7 ottobre 2026, n.179, *Gazzetta Ufficiale della Repubblica Italiana*. Serie generale, n.º 234. Consultado a 8 de julho de 2026. Disponível em: https://www.gazzettaufficiale.it/eli/id/2026/10/08/26G00200/SG

33 ITÁLIA, Decreto legislativo 9 settembre 2026, n. 160, *Gazzetta Ufficiale della Repubblica Italiana*. Serie generale, n.º 214. Consultado a 8 de julho de 2026. Disponível em: https://www.gazzettaufficiale.it/eli/id/2026/09/15/26G00179/sg

34 ESPANHA, Real Decreto n.º 729/2023, de 22 de agosto. *Boletín Oficial del Estado*, 2023. Consultado a 8 de julho de 2026. Disponível em: https://www.boe.es/eli/es/rd/2023/08/22/729/con

atribuições de supervisão, fiscalização e sanção. Meses depois, o *Real Decreto* 817/2023[35], de 8 de novembro, instituiu um ambiente de testagem destinado a ensaiar a proposta de Regulamento. A adaptação ao *RIA* está agora em curso através do *Proyecto de Ley Orgánica para el buen uso y la gobernanza de la nteligência artificial*[36], em apreciação parlamentar. O projeto tem quarenta e um artigos, distribuídos por cinco capítulos: disposições gerais, governação, ambientes de testagem de regulamentação, medidas para o setor público estatal e sanções.

No âmbito da governação, o modelo é descentralizado. A *Dirección General de Inteligencia Artificial* assume as funções de autoridade notificadora (artigo 4.º). A fiscalização do mercado é repartida por várias entidades (artigo 5.º): a AESIA detém a competência geral, a *Agencia Española de Protección de Datos* fica com a recolha não direcionada de imagens faciais e a categorização biométrica, o *Consejo General del Poder Judicial* acompanha os sistemas usados na aplicação da lei e na justiça. O *Banco de España*, a CNMV e a *Dirección General de Seguros* atuam nos respetivos setores. A AESIA é também o ponto de contacto único (artigo 6.º) e incumbe-lhe criar o ambiente de testagem de regulamentação obrigatório do artigo 57.º do RIA, podendo as restantes autoridades criar ambientes adicionais no âmbito das suas atribuições (artigo 11.º).

O regime sancionatório classifica as infrações em muito graves, graves e leves (artigos 13.º a 41.º), com montantes máximos dentro dos limites do *RIA*. O setor público fica sujeito apenas a advertência e a medidas corretivas.

O projeto não contém, no entanto, a habilitação legal que o artigo 5.º, n.º 5, do *RIA* exige. A identificação biométrica à distância em tempo real para aplicação da lei não é, por isso, admitida em Espanha.

Em suma, a diferença que mais releva é que em Espanha ainda não está habilitada a identificação biométrica à distância em tempo real, para efeitos de aplicação da lei, enquanto

35 ESPANHA, Real Decreto n.º 817/2023, de 8 de novembro. *Boletín Oficial del Estado*, 2023. Consultado a 8 de julho de 2026. Disponível em: https://www.boe.es/eli/es/rd/2023/11/08/817

36 ESPANHA, Cortes Generales. Congreso de los Diputados. Proyecto de Ley Orgánica 121/0000096. *Boletín Oficial de las Cortes Generales:* Congreso de los Diputados. Serie A, n.º 97-1, 12 de junho de 2026, pp. 1-41. Consultado em 17 de julho de 2026. Disponível em: https://www.congreso.es/public_oficiales/L15/CONG/BOCG/A/BOCG-15-A-97-1.PDF

em Itália já está consagrada na lei, Espanha ainda não procedeu nesse sentido, pelo que permanece aí proibida.

VI. SÚMULA DAS OPÇÕES MAIS RELEVANTES PERFILHADAS

É procedida agora a uma síntese das principais opções adotadas ao longo da proposta do anteprojeto de lei de execução do RIA, procurando explanar os motivos determinantes para as soluções preconizadas nas matérias em que o Regulamento deixa margem de decisão ao legislador nacional. Esta resenha segue a estrutura dos capítulos que compõem essa proposta.

Capítulo I – Disposições gerais

O primeiro capítulo parte de uma escolha de base em que a proposta não pretende criar uma lei geral sobre IA, como ocorre no caso italiano, mas executar o RIA, intervindo apenas nas matérias em que o próprio Regulamento faculta margem de decisão aos Estados-Membros ou exige medidas nacionais de execução. Esta delimitação procura obviar a criação de um regime paralelo ou sobreposto com matérias já harmonizadas pelo direito da União Europeia. É por essa razão que o elenco de definições que constam no artigo 3.º do RIA não é reproduzido na proposta, apenas são introduzidas as definições conexionadas à governação nacional da IA. A mesma preocupação explica a solução adotada para a interpretação e aplicação da lei. O RIA, os seus atos delegados e de execução e as orientações da Comissão constituem as referências primordiais, ficando as orientações nacionais num plano subordinado. Pretende-se, assim, assegurar uma aplicação coerente do Regulamento e evitar que cada autoridade nacional desenvolva uma interpretação própria.

A utilização do português é estabelecida para a informação destinada aos responsáveis pela implantação e às pessoas afetadas, enquanto, para documentação técnica e declarações de conformidade, se admite a utilização do inglês caso seja aceite pela autoridade competente. Quanto à literacia em IA, a proposta vai além da simples reprodução do RIA e impõe aos serviços e organismos da Administração Pública que utilizem sistemas de IA o dever de assegurar formação adequada aos trabalhadores envolvidos na sua operação e utilização. De facto, a supervisão humana só é efetiva quando quem opera e utiliza os sistemas dispõe dos

conhecimentos suficientes e necessários para compreender, e também questionar, os respetivos resultados.

Capítulo II – Governação e supervisão

A governação constitui uma das escolhas nacionais mais relevantes, porque o RIA estabelece obrigações institucionais, mas não impõe um único modelo administrativo. A opção foi não criar uma nova autoridade especializada e aqui afastamo-nos do modelo espanhol. Atendendo à opção governamental, já veiculada, ainda que não formalmente, de designar a ANACOM como ponto de contacto único e autoridade de fiscalização do mercado, mantivemos essa escolha, que permite beneficiar de uma estrutura regulatória já existente, evitando os custos administrativos associados à criação de uma nova autoridade. Ao mesmo tempo, a proposta não concentra nela toda a fiscalização. Mantêm-se as competências das autoridades setoriais nos domínios em que o próprio RIA determina ou aconselha uma autoridade específica, à semelhança dos modelos espanhol e italiano, tendo a ANACOM uma competência residual, ou seja, é a autoridade de fiscalização do mercado competente nos domínios que não caibam a outra autoridade.

Para uma articulação mais efetiva das autoridades nacionais competentes e mesmo das autoridades responsáveis pela proteção dos direitos fundamentais é proposta a criação do Conselho de Coordenação para a IA, presidido pela ANACOM. Também foi opção não antecipar todas as designações concretas das autoridades. Em matérias cuja repartição depende da organização administrativa e das competências setoriais, a concretização é remetida para decreto-lei, o que permite manter alguma flexibilidade sem alterar a estrutura basilar da lei.

A proposta estabelece, ainda, a criação de um Registo Nacional, conforme exigência imposta pelo RIA, quanto aos sistemas de risco elevado do ponto 2 do Anexo III, relativos a infraestruturas críticas, e, por outro lado, reúne num só sistema notificações, pedidos, avaliações, medidas corretivas e decisões sancionatórias definitivas. Finalmente, atendendo a que o RIA impõe aos Estados-Membros vários deveres de informação à Comissão Europeia, é estabelecido um mecanismo de reporte centralizado. Num modelo plural, a informação dirigida à Comissão Europeia poderia ser produzida de forma dispersa, deste modo a ANACOM recebe a função de coordenar a sua recolha e apresentação.

Capítulo III – Ambiente de testagem da regulamentação da IA e testagem em condições reais

Neste capítulo, a questão foi saber se poderia existir este mecanismo de promoção da inovação prevista no RIA sem criar uma estrutura administrativa nova e a escolha recaiu nas Zonas Livres Tecnológicas, previstas no Decreto-Lei n.º 67/2021l, de 30 de julho. A proposta permite que o ambiente nacional funcione no interior dessas zonas, aplicando-se subsidiariamente o respetivo regime. A opção evita duplicações e aproveita uma estrutura de experimentação já existente. Quanto à participação, a proposta privilegia micro, pequenas e médias empresas, empresas em fase de arranque e projetos de interesse público. A razão é permitir que o acesso à experimentação não fique reservado aos operadores com maior capacidade financeira e técnica. A gratuitidade segue a mesma lógica. A participação no ambiente de testagem não afasta a responsabilidade pelos danos causados a terceiros. O participante apenas fica isento de coimas por infrações cometidas durante a participação, desde que tenha cumprido o plano e atuado de boa-fé.

Na testagem em condições reais, o RIA deixa ao legislador nacional uma escolha relevante quanto à aprovação tácita. A proposta adota-a como regra, sobretudo por razões de celeridade e de atração de investimento. Contudo, afasta-a nos domínios considerados mais sensíveis (biometria, aplicação da lei, migração, asilo e controlo de fronteiras), exigindo autorização expressa. A solução resulta de uma ponderação entre a promoção da inovação e a tutela dos direitos fundamentais.

Capítulo IV – Identificação biométrica à distância para efeitos de aplicação da lei

O Regulamento proíbe a identificação biométrica à distância em tempo real em espaços acessíveis ao público para efeitos de aplicação da lei[37]. Permite, porém, que os Estados-Membros autorizem certas utilizações mediante legislação nacional que fixe as condições. A proposta exerce essa faculdade com regras estreitas, na medida em que está em causa uma tecnologia que afeta a privacidade e a liberdade de circulação de todos os que passam no

37 Sobre esta matéria, vd. DIAS, Vítor Hugo, "*Proibições de identificação biométrica à distância: limites e exceções*". *AI Act - o novo regulamento de inteligência artificial europeu.* Coimbra: Almedina, 2025, pp.243-253; LYNCH, Nessa, Facial Recognition Technology in Policing and Security - Case Studies in Regulation, *Laws* 2024, *13*(3), 35. Consultado a 9 de julho de 2026. Disponível em: https://www.mdpi.com/2075-471X/13/3/35

espaço vigiado, e não apenas de quem é procurado. São admitidos os três objetivos do RIA: busca de vítimas e de pessoas desaparecidas, prevenção de ameaças iminentes para a vida ou de ataques terroristas, e localização de suspeitos de crimes do anexo II puníveis com pena máxima não inferior a quatro anos. A utilização tem de ser estritamente necessária, limitada à confirmação da identidade da pessoa visada e os objetivos têm de ser impossíveis ou muito difíceis de alcançar em tempo útil por outros meios, numa formulação próxima à estabelecida no n.º 1 do artigo 187.º do Código de Processo Penal.

A principal opção é a reserva do juiz, também na utilização em diferido, em detrimento da autoridade administrativa independente, em razão da intensidade da restrição e da usual reserva judicial em matérias de forte ingerência na vida privada. A mesma opção é seguida para a utilização em diferido. O juiz decide em 48 horas e pode autorizar parcialmente. A autorização dura até 72 horas nos casos de ameaça e até 15 dias nos restantes. É prorrogável enquanto os pressupostos subsistirem, e a utilização cessa quando deixem de existir. Em caso de urgência, a utilização pode começar antes, mas a autorização deve ser requerida em 24 horas. Se for recusada, os dados são eliminados e não valem como prova. Cada utilização é proposta à CNPD e à autoridade de fiscalização do mercado, que reportam anualmente à Comissão Europeia, com conhecimento da ANACOM. Na utilização em diferido, a autorização pode ser pedida até 48 horas após o início.

A proposta procura ainda evitar que a identificação biométrica seja convertida numa prova automática[38]. Nenhuma decisão com efeitos jurídicos adversos pode assentar exclusivamente no resultado do sistema. A dupla verificação humana é mantida como regra e as bases de dados utilizadas ficam sujeitas a requisitos de legalidade e rastreabilidade. A utilização de bases construídas mediante recolha aleatória de imagens faciais é expressamente afastada. Os resultados só passam para outro processo se respeitarem a crime do anexo II, forem indispensáveis à prova e houver despacho judicial. A violação dos requisitos geral nulidade.

Capítulo V – Regime sancionatório

A proposta consigna o modelo português do ilícito de mera ordenação social (RGCO), em vez de criar um regime sancionatório autónomo. A escolha permite integrar as infrações

38 Sobre a prova, neste âmbito, *vd.* SILVA, Francisca Oliveira, "*Prova por Reconhecimento e Inteligência Artificial*". Coimbra: Almedina, 2024.

ao RIA na tradição contraordenacional nacional e sujeitar o procedimento às garantias próprias desse sistema. A classificação em contraordenações muito graves, graves e leves, tal como a solução espanhola, procura acompanhar a estrutura do artigo 99.º do RIA. Não são, contudo, criados limites mínimos especiais, porque se entende que a fixação de mínimos poderia dificultar a graduação da sanção segundo a gravidade concreta da infração. Os critérios de determinação da coima remetem para os elementos indicados pelo próprio RIA.

Quanto às entidades públicas, o RIA permite aos Estados-Membros a possibilidade de lhes aplicar coimas, mas a proposta decide não o fazer. A justificação é que a coima paga por um organismo público representa, em última análise, uma transferência de recursos públicos e pode não produzir o efeito dissuasor pretendido, embora haja o risco de gerar assimetrias. Em alternativa, prevê-se a admoestação, a determinação das medidas necessárias à cessação da infração, as sanções acessórias e a eventual responsabilidade disciplinar. Uma solução próxima à preconizada por Espanha.

Outra escolha relevante é a competência judicial. Todas as impugnações das decisões sancionatórias são concentradas no Tribunal da Concorrência, Regulação e Supervisão. A razão apresentada é evitar que uma mesma infração ao RIA seja apreciada por tribunais diferentes apenas porque foi sancionada por autoridades distintas e, assim, favorecer uma jurisprudência uniformizada. Foi preconizada uma solução idêntica à da Lei n.º 12-A/2026, de 15 de abril (Lei de execução do Regulamento dos Serviços Digitais), plasmada no seu artigo 25.º, na escolha do tribunal competente.

Relativamente ao problema da sobreposição entre regimes em que uma mesma conduta pode violar simultaneamente o RIA e outros diplomas, como o RGPD ou o Regulamento dos Serviços Digitais[39], a proposta estabelece mecanismos de coordenação entre autoridades e impede a dupla sanção pelos mesmos factos e fundamento, no respeito pelo princípio *ne bis in idem*.

39 Sobre este assunto, *vd.* BESSA, Tiago, As interseções entre o Regulamento dos Serviços Digitais e o Regulamento da Inteligência Artificial. *AI Act – o novo regulamento de inteligência artificial europeu.* Coimbra: Almedina, 2025, pp.55-78; ARTZT, Matthias, AI & the GDPR: How to Reconcile Both Areas from the European Law Perspective, *International In-house Counsel Journal,* Nov-2022, Consultado em 16 de agosto de 2026. Disponível em: https://www.iicj.net/paper/1655.
Vietnamese Journal of Legal Sciences, Vol. 07, No. 02, 2022, pp. 39-58. Consultado 16 de agosto de 2026. Disponível em:

Quanto ao produto das coimas, a repartição entre Estado e autoridade sancionadora tem uma finalidade específica, que é financiar o reforço da capacidade técnica de supervisão.

Capítulo VI – Disposições transitórias e finais

Neste último capítulo, a primeira opção é de articulação com o direito da proteção de dados. A proposta deixa claro que o RIA e a lei nacional de execução não substituem o RGPD, a legislação nacional de proteção de dados ou as competências da CNPD.

A proposta clarifica, ainda, que o tratamento de categorias especiais de dados pessoais para correção de enviesamentos, previsto no artigo 4.º-A do RIA, se considera autorizado por lei, para efeitos da Lei n.º 59/2019, de 8 de agosto, quando realizado pelas autoridades competentes para efeitos de aplicação da lei, sob fiscalização do CNPD.

As disposições transitórias resultam sobretudo da necessidade de obviar um vazio institucional, logo até à designação definitiva das autoridades responsáveis pela proteção dos direitos fundamentais, mantêm-se as entidades constantes da lista já existente. Paralelamente, é concedido ao Governo um prazo para aprovar os diplomas com a designação das autoridades e à ANACOM um prazo para regulamentar o Registo Nacional e o ambiente de testagem.

Em suma, a proposta foi construída a partir da margem de decisão efetivamente deixada pelo RIA aos Estados-Membros. Na governação, procurou-se conciliar uma autoridade de coordenação – a ANACOM – com a manutenção das competências setoriais. No ambiente de testagem, a prioridade foi aproveitar estruturas nacionais existentes e facilitar a inovação. No domínio da identificação biométrica, a margem europeia foi exercida de forma mais restritiva, através de autorização judicial, limites temporais, controlo humano e regras probatórias, atendendo à tutela dos direitos fundamentais. No regime sancionatório, optou-se pela não aplicação de coimas às entidades públicas e pela concentração da impugnação judicial num tribunal especializado. Finalmente, as disposições transitórias procuram assegurar que estas escolhas possam ser implementadas sem sobreposição com o direito da proteção de dados.

VII. CONSIDERAÇÕES FINAIS

Face ao antedito, o Regulamento Europeu da Inteligência Artificial estabelece um quadro jurídico harmonizado a nível europeu. A sua aplicação efetiva depende, em larga medida, das medidas adotadas por cada Estado-Membro para assegurar a sua execução. Compete, nomeadamente, ao legislador nacional definir o modelo de governação da IA, criar o ambiente de testagem da regulamentação, disciplinar a utilização de sistemas de identificação biométrica à distância para efeitos de aplicação da lei e estabelecer o regime sancionatório aplicável. Estas opções devem ser ajustadas às especificidades do ordenamento jurídico português, equilibrando a promoção da inovação tecnológica com a salvaguarda efetiva dos direitos fundamentais. A proposta segue, por isso, uma abordagem coerente: concretiza as matérias cuja execução é exigida pelo RIA, adota as soluções consideradas mais adequadas nos casos em que este confere margem de decisão aos Estados-Membros e evita introduzir regras nacionais adicionais em áreas já harmonizadas pelo direito da EU.

É neste contexto que se apresenta, de seguida, uma proposta de anteprojeto de lei de execução, como contributo para o debate e para o subsequente processo legislativo nacional.

VIII. PROPOSTA DE ANTEPROJETO DE LEI EXECUÇÃO, NA ORDEM JURÍDICA INTERNA, DO REGULAMENTO (EU) 2024/1689 DO PARLAMENTO EUROPEU E DO CONSELHO, DE 13 DE JUNHO DE 2024 (REGULAMENTO EUROPEU DA INTELIGÊNCIA ARTIFICIAL)

Exposição de motivos

A Inteligência Artificial (IA), pela sua transversalidade, suscita novos desafios à atividade económica, ao funcionamento das empresas e da Administração Pública e, de modo particularmente relevante, à vida em sociedade, sobretudo no domínio da proteção dos direitos fundamentais. A disseminação de sistemas de IA à escala mundial, pelos benefícios que proporciona em matéria de inovação e eficiência, mas também pelos riscos que pode comportar, torna necessário um quadro jurídico adequado. Esse quadro deve assentar em princípios éticos que orientem o desenvolvimento e a utilização da IA e assegurem a supervisão humana, a definição clara de responsabilidades e a instituição de mecanismos de fiscalização adequados.

Com esse intuito, a União Europeia aprovou o Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, que cria regras harmonizadas em matéria de inteligência artificial (Regulamento da Inteligência Artificial - RIA). O RIA institui um quadro jurídico comum da União Europeia assente numa abordagem baseada no risco. Este quadro foi objeto de alteração pelo Regulamento (UE) 2026/1744 do Parlamento Europeu e do Conselho, de 8 de julho de 2026 ("*Omnibus Digital on AI*"), que simplificou as regras harmonizadas e reviu o calendário de cumprimento de determinadas obrigações, em especial as relativas aos sistemas de IA de risco elevado. O mesmo diploma reforçou igualmente a proteção dos direitos fundamentais, ao introduzir novas proibições relativas à geração ou manipulação de conteúdos íntimos não consentidos e de material de abuso sexual de crianças. Embora o RIA seja diretamente aplicável em Portugal, cabe aos Estados-Membros designar as autoridades competentes, criar ambientes de testagem da regulamentação e estabelecer o regime sancionatório. O RIA admite ainda que estes prevejam, por lei nacional, a utilização dos sistemas de identificação biométrica à distância para efeitos de aplicação da lei. O

presente diploma assume, assim, a natureza de lei de execução e limita-se a regular as matérias que o RIA remete para os Estados-Membros. Por esse motivo, não reproduz os princípios em que o RIA assenta, como a supervisão humana, a transparência, a não discriminação e a responsabilização, que decorrem diretamente do Regulamento e orientam a interpretação e a aplicação do presente regime.

Neste quadro, é consagrada a designação da Autoridade Nacional de Comunicações (ANACOM) como ponto de contacto único nacional e autoridade de fiscalização do mercado com competência residual, sem prejuízo das atribuições das autoridades setoriais. Compete-lhe assegurar a articulação com a Comissão Europeia e com os restantes Estados-Membros e promover a coordenação entre as entidades envolvidas na aplicação do RIA. Para o efeito, é instituído o Conselho de Coordenação para a IA, presidido pela ANACOM, destinado a assegurar a coerência da atuação dessas entidades. É ainda criado o Registo Nacional de Sistemas de IA, gerido pela ANACOM, enquanto instrumento de apoio à fiscalização e à partilha de informação.

Para promover a inovação, é instituído um ambiente nacional de testagem da regulamentação da IA, destinado à experimentação de soluções em condições controladas. O acesso a este ambiente é gratuito para as micro, pequenas e médias empresas, incluindo as empresas em fase de arranque, a fim de fomentar o desenvolvimento e a experimentação de novas soluções tecnológicas por estas entidades.

No domínio da aplicação da lei, o regime de utilização de sistemas de identificação biométrica à distância procura equilibrar as exigências de liberdade e de segurança, assegurando a salvaguarda dos direitos fundamentais. A utilização "em tempo real" em espaços acessíveis ao público fica limitada aos objetivos previstos no presente diploma, dentro dos limites do RIA. Depende ainda de autorização judicial prévia, incluindo quando não exista processo penal em curso. Em situação de urgência devidamente justificada, a utilização pode iniciar-se sem autorização. Nesse caso, a autorização deve ser requerida no prazo máximo de 24 horas, e o juiz decide em igual prazo. Se a autorização não for requerida ou for recusada, total ou parcialmente, a utilização cessa de imediato na parte não autorizada. Os dados e os resultados obtidos são eliminados e não podem ser utilizados como meio de prova nem para quaisquer outros fins. A autorização é concedida pelo prazo estritamente necessário, até 72 horas na prevenção de ameaças iminentes ou, na busca de vítimas e na

localização de suspeitos, até 15 dias, prorrogável mediante fundamentação autónoma. Não se fixa um limite global para as prorrogações, que o RIA não impõe e que a natureza das situações abrangidas desaconselha. A garantia reside na reapreciação judicial de cada prorrogação, cessando a utilização logo que deixem de se verificar os respetivos pressupostos. A utilização em diferido para a pesquisa seletiva de pessoas suspeitas ou condenadas fica igualmente sujeita a controlo judicial. As utilizações não seletivas ou aleatórias são proibidas, em salvaguarda de direitos constitucionalmente consagrados. Exige-se, em qualquer caso, a verificação humana das correspondências obtidas, e nenhuma decisão pode ser tomada exclusivamente com base no resultado do sistema.

Em matéria sancionatória, estabelece-se um regime contraordenacional que distingue infrações muito graves, graves e leves, em conformidade com os limites e critérios previstos no RIA. A impugnação judicial das decisões sancionatórias compete ao Tribunal da Concorrência, Regulação e Supervisão, e aplica-se subsidiariamente o regime geral do ilícito de mera ordenação social. Ficam ressalvadas as competências atribuídas às diferentes autoridades e as garantias dos visados.

Por outro lado, e a bem da coerência do sistema, a aplicação do RIA articula-se com os regimes relativos à proteção de dados pessoais e aos serviços digitais. Prevê-se a cooperação entre as autoridades competentes e proíbe-se a dupla sanção pelos mesmos factos. Clarifica-se ainda a base legal para o tratamento de categorias especiais de dados pessoais pelas autoridades competentes para efeitos de aplicação da lei, quando necessário à deteção e correção de enviesamentos, nos termos e com as garantias previstas no RIA.

Em suma, o presente regime visa assegurar a aplicação efetiva do RIA em Portugal, conciliando a inovação e o desenvolvimento tecnológico com a proteção dos direitos fundamentais, os princípios éticos e a confiança dos cidadãos e das organizações.

CAPÍTULO I
Disposições gerais

Artigo 1.º
Objeto

1 - A presente lei assegura a execução, na ordem jurídica interna, do Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, que cria regras harmonizadas em matéria de inteligência artificial (IA) e que altera os Regulamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e as Diretivas 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828, na sua redação atual, doravante designado por Regulamento da Inteligência Artificial (RIA).

2 - Para efeitos do disposto no número anterior, a presente lei estabelece, nomeadamente:

a) A designação das autoridades nacionais competentes e os mecanismos de coordenação entre elas;
b) O Registo Nacional de Sistemas de IA;
c) O regime dos ambientes de testagem da regulamentação da IA e o da testagem em condições reais;
d) O regime de autorização da utilização, para efeitos de aplicação da lei, de sistemas de identificação biométrica à distância "em tempo real" em espaços acessíveis ao público e de sistemas de identificação biométrica à distância em diferido;
e) O regime sancionatório aplicável às infrações ao RIA.

Artigo 2.º
Âmbito de aplicação

1 - A presente lei aplica-se às pessoas singulares ou coletivas e às demais entidades, públicas ou privadas, abrangidas pelo artigo 2.º do RIA, bem como às autoridades nacionais competentes, no exercício das respetivas funções nesse âmbito.

2 - Excluem-se do âmbito de aplicação da presente lei os sistemas de IA e as atividades excluídos pelo artigo 2.º do RIA, designadamente os sistemas de IA colocados no mercado, colocados em serviço ou utilizados, com ou sem modificações, exclusivamente para fins militares, de defesa ou de segurança nacional, independentemente do tipo de entidade que realize essas atividades.

3 - O regime previsto na presente lei não prejudica a aplicação do Regulamento (UE) 2022/2065 do Parlamento Europeu e do Conselho, de 19 de outubro de 2022, da Lei n.º 12-A/2026, de 15 de abril, e do Decreto-Lei n.º 7/2004, de 7 de janeiro, na sua redação atual, em especial das disposições relativas à responsabilidade dos prestadores de serviços intermediários, em conformidade com o disposto no n.º 5 do artigo 2.º do RIA.

Artigo 3.º
Definições

Para efeitos da presente lei, são aplicáveis as definições constantes do RIA, entendendo-se ainda por:

a) "Autoridade notificadora", a autoridade designada nos termos do artigo 8.º;
b) "Autoridades de fiscalização do mercado", as autoridades designadas nos termos do artigo 9.º;
c) "Autoridades responsáveis pela proteção dos direitos fundamentais", as autoridades e os organismos públicos identificados nos termos do artigo 11.º;
d) "Ponto de contacto único", a autoridade de fiscalização do mercado designada nos termos do artigo 7.º para exercer essa função;
e) "Autoridades nacionais competentes", a autoridade notificadora e as autoridades de fiscalização do mercado, nos termos do ponto 48) do artigo 3.º e do n.º 1 do artigo 70.º do RIA.

Artigo 4.º
Interpretação e aplicação

As disposições da presente lei são interpretadas e aplicadas em conformidade com o RIA e com os atos delegados e de execução adotados ao seu abrigo, tendo em conta as orientações emitidas pela Comissão Europeia sobre a sua aplicação e, em conformidade com estas, as orientações emitidas pela Autoridade Nacional de Comunicações (ANACOM).

Artigo 5.º
Língua

1 - São redigidas em língua portuguesa as instruções de utilização, as informações a prestar aos responsáveis pela implantação e às pessoas afetadas e as demais informações previstas no RIA destinadas a utilizadores em território nacional.
2 - A declaração UE de conformidade e a documentação técnica são disponibilizadas às autoridades nacionais competentes em língua portuguesa, podendo a autoridade aceitar a documentação redigida em língua inglesa.

Artigo 6.º
Literacia em IA

1 - Compete ao Governo, em articulação com a ANACOM e com as demais autoridades nacionais competentes, apoiar e facilitar as medidas de literacia em IA adotadas pelos prestadores e pelos responsáveis pela implantação, com especial atenção às micro, pequenas e médias empresas, nos termos do artigo 4.º do RIA.
2 - Os serviços e organismos da Administração Pública que utilizem sistemas de IA asseguram formação adequada aos trabalhadores envolvidos na sua operação e utilização.
3 - O disposto no presente artigo não prejudica as obrigações que o RIA imponha aos prestadores e aos responsáveis pela implantação.

CAPÍTULO II
Governação e supervisão

SECÇÃO I
Autoridades nacionais competentes

Artigo 7.º
Ponto de contacto único e coordenação

1 - A ANACOM é o ponto de contacto único previsto no n.º 2 do artigo 70.º do RIA e assegura a coordenação das autoridades nacionais competentes e das autoridades responsáveis pela proteção dos direitos fundamentais.

2 - Enquanto ponto de contacto único, incumbe à ANACOM, designadamente:

a) Assegurar a comunicação com a Comissão Europeia e com os pontos de contacto únicos dos demais Estados-Membros;
b) Receber as queixas apresentadas nos termos do artigo 85.º do RIA e dar-lhes seguimento nos termos do artigo 15.º;
c) Consolidar e transmitir à Comissão Europeia a informação prevista no artigo 17.º;
d) Gerir o Registo Nacional de Sistemas de IA.

3 - No exercício das funções de coordenação, incumbe à ANACOM:

a) Emitir orientações destinadas a assegurar a aplicação uniforme do RIA, nos termos do n.º 8 do artigo 70.º do mesmo Regulamento;
b) Promover a cooperação e a troca de informação entre as autoridades referidas no n.º 1;
c) Presidir ao Conselho de Coordenação para a IA;
d) Prestar apoio técnico às autoridades referidas no n.º 1 que o solicitem.

4 - As autoridades nacionais competentes têm em conta as orientações previstas na alínea a) do número anterior e fundamentam expressamente as decisões tomadas em sentido diverso.

5 - O disposto nos números anteriores não prejudica a independência ou a autonomia de que gozem, nos termos da Constituição e da lei, as autoridades referidas no n.º 1, nem as atribuições e os poderes que lhes sejam conferidos.

Artigo 8.º
Autoridade notificadora

1 - A autoridade notificadora é designada por decreto-lei.

2 - Incumbe à autoridade notificadora estabelecer e executar os procedimentos de avaliação, designação e notificação dos organismos de avaliação da conformidade e fiscalizar os organismos notificados, nos termos dos artigos 28.º a 39.º do RIA.

Artigo 9.º
Autoridades de fiscalização do mercado

1 - São autoridades de fiscalização do mercado a ANACOM e as entidades que, no âmbito das respetivas atribuições, sejam designadas para o efeito por decreto-lei.

2 - A ANACOM é a autoridade de fiscalização do mercado competente nos domínios que não caibam a outra autoridade nos termos dos números seguintes ou do decreto-lei referido no número anterior.

3 - As autoridades designadas ao abrigo da legislação de harmonização da União Europeia enumerada na secção A do anexo I do RIA são as autoridades de fiscalização do mercado dos sistemas de IA de risco elevado relacionados com os produtos abrangidos por essa legislação, nos termos do n.º 3 do artigo 74.º do mesmo Regulamento.

4 - As autoridades responsáveis pela supervisão das instituições financeiras sujeitas à legislação da União Europeia em matéria de serviços financeiros são, no âmbito das respetivas atribuições, as autoridades de fiscalização do mercado dos sistemas de IA colocados no mercado, colocados em serviço ou utilizados por essas instituições, nos termos do n.º 6 do artigo 74.º do RIA.

5 - Os sistemas de IA de risco elevado utilizados para efeitos de aplicação da lei, de migração, asilo e controlo de fronteiras, de administração da justiça e de processos democráticos, incluindo os sistemas biométricos utilizados para esses fins, ficam sujeitos à fiscalização da autoridade designada pelo decreto-lei referido no n.º 1, nos termos do n.º 8 do artigo 74.º do RIA.

6 - A autoridade referida no número anterior é designada de entre a Comissão Nacional de Proteção de Dados (CNPD) e as demais autoridades que preencham as condições previstas nos artigos 41.º a 44.º da Diretiva (UE) 2016/680 do Parlamento Europeu e do Conselho, de 27 de abril de 2016.

7 - A fiscalização prevista no n.º 5 não afeta a independência dos tribunais nem interfere com a sua atividade no exercício da função jurisdicional.

8 - As autoridades de fiscalização do mercado exercem, nos respetivos domínios, as atribuições e os poderes previstos no artigo 74.º do RIA e no artigo 14.º do Regulamento (UE) 2019/1020 do Parlamento Europeu e do Conselho, de 20 de junho de 2019, incumbindo-lhes, designadamente:

a) Fiscalizar o cumprimento do RIA;
b) Solicitar aos operadores a documentação, os dados e as informações necessários ao exercício das suas funções;
c) Proceder a investigações e auditorias, por iniciativa própria ou na sequência de queixa;
d) Realizar inspeções, incluindo sem aviso prévio, às instalações que os operadores utilizem para fins profissionais, com exceção das que constituam domicílio;
e) Adotar as medidas corretivas e restritivas previstas nos artigos 79.º a 83.º do RIA;
f) Instaurar e instruir os processos de contraordenação e aplicar as coimas e as sanções acessórias nos termos do capítulo V;

g) Cooperar com as demais autoridades nacionais competentes, nos termos do artigo 12.º, bem como com as autoridades de fiscalização do mercado dos demais Estados-Membros e com a Comissão Europeia;
h) Remeter à ANACOM a informação prevista no artigo 17.º.

9 - Os operadores têm o dever de colaborar com as autoridades de fiscalização do mercado, facultando-lhes o acesso às instalações, à documentação e à informação solicitadas.

Artigo 10.º
Independência, recursos e cibersegurança

1 - As autoridades nacionais competentes exercem as suas funções com independência, imparcialidade e isenção, de modo a salvaguardar a objetividade das suas atividades, nos termos do n.º 1 do artigo 70.º do RIA.

2 - O Estado assegura que as autoridades referidas no número anterior dispõem dos recursos técnicos, financeiros e humanos adequados, bem como das infraestruturas necessárias, ao exercício das suas funções, nos termos do n.º 3 do artigo 70.º do RIA.

3 - Incumbe às autoridades referidas no n.º 1 adotar as medidas necessárias para garantir um nível adequado de cibersegurança, nos termos do n.º 4 do artigo 70.º do RIA.

SECÇÃO II
Autoridades responsáveis pela proteção dos direitos fundamentais

Artigo 11.º
Autoridades responsáveis pela proteção dos direitos fundamentais

1 - São autoridades responsáveis pela proteção dos direitos fundamentais as autoridades e os organismos públicos identificados por resolução do Conselho de Ministros.

2 - Incumbe à ANACOM publicar no seu sítio na Internet a lista das autoridades referidas no número anterior, mantê-la atualizada e notificá-la à Comissão Europeia.

3 - As autoridades responsáveis pela proteção dos direitos fundamentais exercem os poderes previstos no artigo 77.º do RIA em articulação com a autoridade de fiscalização do mercado competente, à qual incumbe organizar os testes que aquelas solicitem.

SECÇÃO III
Cooperação institucional

Artigo 12.º
Dever de cooperação

1 - As autoridades nacionais competentes e as autoridades responsáveis pela proteção dos direitos fundamentais cooperam entre si e com a ANACOM, trocando a informação necessária ao prosseguimento das respetivas atribuições no âmbito do RIA e da presente lei.

2 – O dever de confidencialidade previsto no artigo 78.º do RIA aplica-se a todas as entidades que intervenham na aplicação da presente lei.

3 - A cooperação não prejudica a independência ou a autonomia de que gozem as autoridades referidas no n.º 1, nem as competências das autoridades judiciárias e dos órgãos de polícia criminal no âmbito do processo penal.
4 - As autoridades de fiscalização do mercado articulam a sua atuação com as autoridades competentes ao abrigo do regime dos serviços digitais quando estejam em causa sistemas de IA integrados em serviços intermediários ou disponibilizados através destes, trocando entre si as informações necessárias ao exercício das respetivas atribuições.

Artigo 13.º
Conselho de Coordenação para a IA

1 - É criado o Conselho de Coordenação para a IA, presidido pela ANACOM e composto por representantes das autoridades nacionais competentes.
2 - As autoridades responsáveis pela proteção dos direitos fundamentais participam nas reuniões sempre que estejam em causa matérias das suas atribuições.
3 - Incumbe ao Conselho de Coordenação:

a) Assegurar a coerência da atuação das autoridades que o integram;
b) Emitir parecer sobre os projetos de orientações previstos na alínea a) do n.º 3 do artigo 7.º;
c) Pronunciar-se sobre os conflitos de atribuições e de competência, nos termos do artigo seguinte;
d) Promover ações conjuntas de fiscalização e a partilha de informação, boas práticas e metodologias entre as autoridades que o integram;
e) Contribuir para a estratégia nacional de fiscalização do mercado prevista no artigo 13.º do Regulamento (UE) 2019/1020, de 20 de junho de 2019, no que respeita aos sistemas de IA;
f) Participar na preparação das posições nacionais no Comité Europeu para a IA;
g) Emitir parecer sobre o relatório anual previsto no n.º 5 do artigo 17.º.

4 - O Conselho de Coordenação reúne pela primeira vez no prazo de 60 dias a contar da data de entrada em vigor do decreto-lei previsto nos artigos 8.º e 9.º.
5 - O regulamento interno do Conselho de Coordenação é aprovado nos 60 dias subsequentes à sua primeira reunião.

Artigo 14.º
Conflitos de atribuições e de competência

1 - Quando mais do que uma autoridade se considere competente ou incompetente para a mesma matéria, a questão é submetida ao Conselho de Coordenação.
2 - Se o conflito não for dirimido no prazo de 30 dias a contar da sua submissão ao Conselho de Coordenação, aplica-se o disposto nos artigos 51.º e 52.º do Código do Procedimento Administrativo.
3 - Até à resolução do conflito, a ANACOM pode adotar as medidas provisórias urgentes necessárias para prevenir riscos graves para a saúde, a segurança ou os direitos fundamentais.

SECÇÃO IV
Queixas, Registo Nacional e reporte

Artigo 15.º
Queixas e denúncias

1 - As queixas previstas no artigo 85.º do RIA podem ser apresentadas à ANACOM ou à autoridade de fiscalização do mercado competente.
2 - A autoridade que receba queixa relativa a matéria da competência de outra autoridade remete-a a esta no prazo de 10 dias úteis e informa o queixoso.
3 - À denúncia de infrações ao RIA e à proteção das pessoas que as denunciem é aplicável a Lei n.º 93/2021, de 20 de dezembro, nos termos do artigo 87.º do RIA.
4 - A ANACOM e as autoridades de fiscalização do mercado são, no âmbito das respetivas atribuições, autoridades competentes para receber e dar seguimento a denúncias externas, nos termos do n.º 1 do artigo 12.º da Lei n.º 93/2021, de 20 de dezembro.

Artigo 16.º
Registo Nacional de Sistemas de IA

1 - É criado o Registo Nacional de Sistemas de IA, gerido pela ANACOM, que funciona como sistema de informação de apoio à coordenação e à fiscalização do mercado de sistemas de IA em território nacional.
2 - Os prestadores ou, quando aplicável, os respetivos mandatários inscrevem no Registo Nacional, antes da colocação no mercado ou da colocação em serviço, os sistemas de IA de risco elevado referidos no ponto 2 do anexo III do RIA, nos termos do n.º 5 do artigo 49.º do mesmo Regulamento.
3 - Para além das inscrições previstas no número anterior, e sem que daí resultem obrigações de inscrição para os operadores, do Registo Nacional constam:

a) As notificações, requerimentos e comunicações dirigidos às autoridades de fiscalização do mercado, designadamente os relativos à utilização de sistemas de identificação biométrica à distância "em tempo real", às autorizações excecionais de colocação no mercado ou em serviço, aos incidentes graves, à testagem em condições reais e aos ambientes de testagem da regulamentação da IA;
b) As avaliações de não conformidade e de risco e as medidas corretivas e restritivas adotadas pelas autoridades de fiscalização do mercado, nos termos dos artigos 79.º a 83.º do RIA;
c) As decisões condenatórias proferidas em processo de contraordenação que se tenham tornado definitivas ou transitado em julgado.

4 - Não são inscritos no Registo Nacional os modelos de IA de finalidade geral, cuja supervisão compete à Comissão Europeia, sem prejuízo da inclusão da informação sobre esses modelos que esta transmita às autoridades de fiscalização do mercado.
5 - As autoridades nacionais competentes têm acesso direto ao Registo Nacional, na estrita medida do necessário ao prosseguimento das respetivas atribuições.

6 - A informação relativa aos sistemas referidos no n.º 5 do artigo 9.º é mantida em secção de acesso restrito, reservada à ANACOM e à autoridade de fiscalização do mercado competente para esses sistemas.
7 - O tratamento de dados pessoais no Registo Nacional rege-se pela legislação aplicável em matéria de proteção de dados pessoais, ficando as entidades com acesso ao Registo Nacional sujeitas a dever de sigilo quanto à informação a que acedam.
8 - A ANACOM aprova, por regulamento, ouvidas as autoridades de fiscalização do mercado e a CNPD, as regras de organização e funcionamento do Registo Nacional, designadamente os perfis de acesso e os prazos de conservação da informação.
9 - O Registo Nacional não substitui a base de dados da União Europeia (UE) prevista no artigo 71.º do RIA.

Artigo 17.º
Reporte

1 - As autoridades nacionais competentes remetem anualmente à ANACOM, no âmbito das respetivas atribuições e em prazo fixado pelo Conselho de Coordenação, a informação relativa ao ano civil anterior necessária ao cumprimento dos deveres de reporte previstos no RIA, designadamente sobre:

a) As atividades de fiscalização do mercado com potencial interesse para a aplicação das regras de concorrência da União Europeia, a utilização de práticas proibidas e as medidas adotadas, nos termos do n.º 2 do artigo 74.º do RIA;
b) As coimas aplicadas, bem como os litígios ou processos judiciais conexos, nos termos do n.º 11 do artigo 99.º do RIA;
c) Os ambientes de testagem da regulamentação da IA, nos termos do n.º 16 do artigo 57.º do RIA.

2 - As autoridades nacionais competentes remetem ainda à ANACOM, com a periodicidade prevista no n.º 6 do artigo 70.º do RIA e em prazo fixado pelo Conselho de Coordenação, a informação sobre os recursos financeiros e humanos de que dispõem, incluindo uma avaliação da sua adequação.
3 - As autoridades de fiscalização do mercado dão conhecimento à ANACOM, em simultâneo, das comunicações que efetuem à Comissão Europeia, aos demais Estados-Membros ou ao Banco Central Europeu, ao abrigo do n.º 11 do artigo 73.º, do segundo parágrafo do n.º 7 do artigo 74.º e dos artigos 79.º a 83.º do RIA.
4 - Incumbe à ANACOM coordenar, sem prejuízo dos deveres de comunicação diretamente previstos no RIA para as autoridades nacionais competentes, a recolha e a apresentação à Comissão Europeia da informação cuja transmissão seja da responsabilidade do Estado Português.
5 - A ANACOM elabora e publica no seu sítio na Internet um relatório anual sobre a aplicação do RIA em Portugal, sem prejuízo das obrigações de confidencialidade previstas no artigo 78.º do RIA e das normas aplicáveis à proteção de dados pessoais.

6 - Para efeitos do n.º 5 do artigo 5.º do RIA, a ANACOM notifica à Comissão Europeia as regras previstas na secção I do capítulo IV da presente lei, no prazo de 30 dias após a sua publicação.

CAPÍTULO III
Ambiente de testagem da regulamentação da IA e testagem em condições reais

Artigo 18.º
Criação, finalidade e supervisão

1 - É criado o ambiente nacional de testagem da regulamentação da IA (*"sandbox regulatória"*), doravante "ambiente de testagem", previsto no artigo 57.º do RIA.
2 - O ambiente de testagem destina-se ao desenvolvimento, treino, testagem e validação de sistemas de IA inovadores, em ambiente controlado e por período limitado, antes da sua colocação no mercado ou colocação em serviço.
3 - Cada participação obedece a um plano específico acordado entre o potencial prestador e a ANACOM.
4 - O ambiente de testagem funciona sob a supervisão da ANACOM, em articulação com:
a) A autoridade de fiscalização do mercado competente em razão da matéria, quando não seja a ANACOM;
b) A CNPD, sempre que os sistemas testados envolvam o tratamento de dados pessoais, nos termos do n.º 10 do artigo 57.º do RIA.

5 - O ambiente de testagem pode incluir testagem em condições reais, nos termos do n.º 4 do artigo 58.º e do artigo 60.º do RIA.

Artigo 19.º
Articulação com as Zonas Livres Tecnológicas

Quando o ambiente de testagem funcione no âmbito de uma Zona Livre Tecnológica, é subsidiariamente aplicável, com as necessárias adaptações, o Decreto-Lei n.º 67/2021, de 30 de julho.

Artigo 20.º
Acesso e critérios de seleção

1 - O acesso ao ambiente de testagem faz-se mediante candidatura dirigida à ANACOM.
2 - Têm prioridade:
a) As micro, pequenas e médias empresas, incluindo as empresas em fase de arranque, com sede ou estabelecimento estável em Portugal;
b) Os projetos com aplicação em domínios de interesse público.

3 - O acesso é gratuito para as entidades referidas na alínea a) do número anterior, sem prejuízo da recuperação de custos excecionais nos termos da alínea d) do n.º 2 do artigo 58.º do RIA.

4 - Os critérios de elegibilidade e de seleção, o número de vagas por ciclo e a duração máxima de participação são fixados por regulamento da ANACOM, ouvidas as autoridades de fiscalização do mercado designadas nos termos do artigo 9.º e a CNPD, com observância dos atos de execução adotados ao abrigo do n.º 1 do artigo 58.º do RIA.
5 - As referências às micro, pequenas e médias empresas constantes do presente artigo abrangem as pequenas empresas de média capitalização, na medida em que o RIA lhes estenda o regime aplicável às micro, pequenas e médias empresas.

Artigo 21.º
Efeitos da participação

1 - A participação no ambiente de testagem não isenta o participante do cumprimento das obrigações previstas no RIA nem da responsabilidade pelos danos causados a terceiros.
2 - Ao participante que respeite o plano específico e os termos e condições da participação e siga de boa-fé as orientações da ANACOM não são aplicadas coimas por infrações ao RIA cometidas no decurso da participação, nos termos do n.º 12 do artigo 57.º do RIA.
3 - A ANACOM pode suspender, temporária ou definitivamente, a testagem sempre que identifique riscos significativos para a saúde, a segurança ou os direitos fundamentais que não possam ser eficazmente atenuados.
4 - Findo cada ciclo, a ANACOM publica um relatório com as principais conclusões do ambiente de testagem, salvaguardados os segredos comerciais e os dados pessoais.

Artigo 22.º
Testagem em condições reais

1 - O pedido de aprovação da testagem em condições reais, nos termos da alínea b) do n.º 4 do artigo 60.º do RIA, é apresentado, acompanhado do plano de testagem em condições reais, à autoridade de fiscalização do mercado competente em razão da matéria ou, na falta desta, à ANACOM.
2 - A autoridade decide no prazo de 30 dias a contar da receção do pedido, considerando-se a testagem e o plano tacitamente aprovados decorrido esse prazo sem decisão.
3 - O disposto na parte final do número anterior não se aplica aos sistemas de IA de risco elevado nos domínios da biometria, da aplicação da lei e da migração, asilo e gestão do controlo das fronteiras, mencionados, respetivamente, nos pontos 1, 6 e 7 do anexo III do RIA, cuja testagem depende de autorização expressa, a proferir no prazo previsto no n.º 2.
4 - O prazo previsto no n.º 2 suspende-se enquanto o procedimento estiver parado por motivo imputável ao requerente, designadamente na pendência da apresentação de elementos adicionais solicitados pela autoridade competente.
5 - A autoridade competente dá conhecimento da decisão à ANACOM e, quando a testagem envolva o tratamento de dados pessoais, à CNPD.
6 - A autoridade competente exerce, quanto à testagem em curso, os poderes previstos no artigo 76.º do RIA, podendo exigir alterações, suspender a testagem ou fazê-la cessar.

CAPÍTULO IV
Identificação biométrica à distância para efeitos de aplicação da lei

SECÇÃO I
Utilização "em tempo real"

Artigo 23.º
Admissibilidade e objetivos

1 - A utilização de sistemas de identificação biométrica à distância "em tempo real" em espaços acessíveis ao público, para efeitos de aplicação da lei, só é admitida para os objetivos previstos no número seguinte e nas condições estabelecidas na alínea h) do n.º 1 e nos n.ºs 2 a 6 do artigo 5.º do RIA e nos artigos 24.º a 28.º da presente lei.

2 - Para efeitos do n.º 5 do artigo 5.º do RIA, a utilização pode ser autorizada para a prossecução dos seguintes objetivos:

a) A busca seletiva de vítimas específicas de rapto, de tráfico de seres humanos ou de exploração sexual de seres humanos, bem como a busca de pessoas desaparecidas;
b) A prevenção de uma ameaça específica, substancial e iminente para a vida ou a segurança física de pessoas singulares, ou de uma ameaça real e atual ou real e previsível de ataque terrorista;
c) A localização ou identificação de pessoa suspeita da prática de crime correspondente a uma das infrações penais referidas no anexo II do RIA, punível com pena de prisão de duração máxima não inferior a quatro anos, para efeitos de investigação criminal, de instauração da ação penal ou de execução de pena.

Artigo 24.º
Pressupostos materiais

1 - A utilização só pode ser autorizada quando seja estritamente necessária para um dos objetivos previstos no n.º 2 do artigo anterior, os mesmos sejam impossíveis ou muito difíceis de alcançar em tempo útil por outros meios e se destine apenas a confirmar a identidade da pessoa ou das pessoas especificamente visadas.

2 - Na decisão são ponderadas:

a) A natureza da situação que dá origem à eventual utilização, em especial a gravidade, a probabilidade e a escala dos danos que seriam causados caso o sistema não fosse utilizado;
b) As consequências da utilização para os direitos e liberdades de todas as pessoas afetadas, em especial a gravidade, a probabilidade e a escala dessas consequências.

3 - A utilização fica sujeita às limitações temporais, geográficas e pessoais fixadas na autorização.

Artigo 25.º
Autorização judicial

1 - A utilização de sistemas de identificação biométrica à distância "em tempo real" para os objetivos previstos no n.º 2 do artigo 23.º depende de autorização judicial, requerida pela autoridade responsável pela aplicação da lei competente, nos termos da lei processual penal.
2 - Se a utilização não se referir a processo penal em curso, a autorização é requerida pela autoridade de polícia criminal competente ao juiz de instrução criminal da sede da entidade competente para a investigação criminal.
3 - O requerimento indica:
a) O objetivo prosseguido e, nos casos previstos na alínea c) do n.º 2 do artigo 23.º, o crime em causa e a infração penal do anexo II do RIA a que corresponde;
b) A pessoa ou as pessoas visadas ou os elementos que permitam a sua identificação;
c) O âmbito territorial e o período da utilização;
d) Os elementos objetivos ou os indícios claros que fundamentam a necessidade da utilização;
e) A identificação da base de dados.
4 - A avaliação de impacto sobre os direitos fundamentais prevista no artigo 27.º do RIA e o registo na base de dados da UE previsto no artigo 49.º do RIA são realizados antes da primeira utilização de cada sistema de identificação biométrica à distância "em tempo real" e atualizados quando necessário, bastando ao requerimento a respetiva referência, sem prejuízo do disposto no n.º 3 do artigo seguinte quanto ao registo.
5 - O juiz decide no mais curto prazo, que não pode exceder 48 horas, podendo autorizar a utilização total ou parcialmente.
6 - O despacho é fundamentado, fixando o objetivo prosseguido, a pessoa ou as pessoas visadas, o âmbito territorial, o prazo da utilização e a identificação da base de dados.
7 - O processo de autorização tem natureza urgente, sem prejuízo do regime do segredo de justiça.

Artigo 26.º
Procedimento de urgência

1 - Em situação de urgência devidamente justificada, a utilização pode iniciar-se sem autorização judicial prévia, por determinação da autoridade responsável pela aplicação da lei competente.
2 - A autorização é requerida nos termos do artigo anterior, sem demora injustificada e, no máximo, no prazo de 24 horas após o início da utilização, devendo o juiz decidir em igual prazo.
3 - O registo do sistema de identificação biométrica à distância "em tempo real" na base de dados da UE pode ser concluído após o início da utilização, no mais curto prazo.
4 - Se a autorização não for requerida no prazo previsto no n.º 2 ou for recusada, a utilização cessa de imediato e todos os dados e resultados dela decorrentes são eliminados, não podendo ser utilizados como meio de prova nem para quaisquer outros fins.

5 - Se a autorização for concedida parcialmente, a utilização cessa de imediato na parte não autorizada, aplicando-se o disposto no número anterior aos dados e resultados obtidos fora dos limites fixados no despacho.

Artigo 27.º
Duração da autorização

1 - A autorização é concedida pelo prazo estritamente necessário, não superior a 72 horas nos casos previstos na alínea b) do n.º 2 do artigo 23.º, e a 15 dias nos casos previstos nas alíneas a) e c), prorrogável por iguais períodos mediante requerimento fundamentado que demonstre a subsistência dos pressupostos.
2 - A utilização do sistema cessa logo que deixem de se verificar os pressupostos que fundamentaram a autorização, ainda que o prazo autorizado não se encontre esgotado.

Artigo 28.º
Comunicação e relatórios

1 - Cada utilização é comunicada pela entidade utilizadora, no prazo de 10 dias úteis após a sua cessação, à CNPD e, quando distinta, à autoridade de fiscalização do mercado competente, com exclusão dos dados operacionais sensíveis.
2 - As autoridades referidas no número anterior apresentam à Comissão Europeia os relatórios anuais previstos no n.º 6 do artigo 5.º do RIA.
3 – As autoridades mencionadas no n.º 1 dão conhecimento à ANACOM, em simultâneo, dos relatórios apresentados nos termos do número anterior, com exclusão dos dados operacionais sensíveis.

SECÇÃO II
Utilização em diferido

Artigo 29.º
Autorização judicial

1 - A utilização de sistemas de identificação biométrica à distância em diferido para a pesquisa seletiva de pessoa ou pessoas suspeitas da prática de crime ou condenadas pela sua prática depende de autorização judicial, nos termos do n.º 10 do artigo 26.º do RIA.
2 - A autoridade responsável pela aplicação da lei competente requer a autorização ao juiz, previamente ou sem demora injustificada e, no máximo, no prazo de 48 horas após o início da utilização.
3 - O requerimento indica o crime e o processo, a pessoa ou as pessoas visadas ou os elementos que permitam a sua identificação, os motivos da sua utilização e a identificação da base de dados.
4 - O juiz decide no prazo máximo de 48 horas.
5 - Se a autorização for recusada, a utilização cessa de imediato e todos os dados pessoais e resultados dela decorrentes são eliminados, não podendo ser utilizados como meio de prova.

6 - Não depende de autorização a utilização para a identificação inicial de um potencial suspeito com base em factos objetivos e verificáveis diretamente relacionados com o crime, a qual é registada no processo.

Artigo 30.º
Limites e controlo

1 - Cada utilização limita-se ao estritamente necessário para a investigação de um crime concreto.
2 - Não é admitida a utilização não seletiva, sem ligação a um crime, a um processo penal, a uma ameaça real e atual ou real e previsível da prática de um crime, ou à busca de uma pessoa desaparecida específica.
3 - Cada utilização é documentada no processo, sendo o registo disponibilizado, a pedido, à CNPD e, quando distinta, à autoridade de fiscalização do mercado competente, com exclusão dos dados operacionais sensíveis.
4 - Os responsáveis pela implantação apresentam anualmente às autoridades referidas no número anterior um relatório agregado sobre as utilizações efetuadas, com exclusão dos dados operacionais sensíveis.

SECÇÃO III
Disposições comuns

Artigo 31.º
Garantias comuns

1 - Nenhuma decisão que produza efeitos jurídicos adversos para uma pessoa pode ser tomada exclusivamente com base no resultado de um sistema de identificação biométrica à distância.
2 - Os dados biométricos captados ou recolhidos através dos sistemas previstos no presente capítulo só podem ser comparados com dados constantes de bases de dados de referência legalmente constituídas para efeitos de prevenção, deteção, investigação ou repressão de infrações penais ou de execução de sanções penais, ou a que a autoridade responsável pela aplicação da lei competente tenha acesso nos termos da lei para esses efeitos, ou que tenham sido legitimamente obtidos no âmbito do processo.
3 - Não podem ser utilizadas como bases de dados de referência as bases de dados de reconhecimento facial criadas ou ampliadas através de sistemas proibidos nos termos da alínea e) do n.º 1 do artigo 5.º do RIA.
4 - Nenhuma ação ou decisão é tomada com base na identificação resultante do sistema sem que esta tenha sido verificada e confirmada separadamente por, pelo menos, duas pessoas singulares com a competência, a formação e a autoridade necessárias, salvo quando, em situação de urgência, tal se revele desproporcionado, nos termos do n.º 5 do artigo 14.º do RIA.

5 - Ao abrigo do n.º 5 do artigo 14.º do RIA, considera-se desproporcionada a exigência prevista no número anterior em situação de urgência ou de utilização em tempo real em que a intervenção de uma segunda pessoa não seja possível em tempo útil. Nesses casos, a verificação é efetuada por uma pessoa com competência, a formação e a autoridade necessárias, e é confirmada por uma segunda pessoa no mais curto prazo.

Artigo 32.º
Utilização dos resultados

1 - Os resultados obtidos nos termos do presente capítulo podem ser utilizados no processo em que a utilização foi autorizada ou determinada.
2 - Os resultados podem ser utilizados noutro processo, em curso ou a instaurar, quando respeitem a crime referido na alínea c) do n.º 2 do artigo 23.º e se mostrem indispensáveis à prova, mediante despacho do juiz do processo em que devam ser utilizados.
3 - São nulos, não podendo ser utilizados, os resultados obtidos com violação dos requisitos e condições previstos nos artigos 23.º a 27.º e 29.º, nos n.ºs 1 e 2 do artigo 30.º e no artigo 31.º.

Artigo 33.º
Proteção de dados pessoais e direito subsidiário

1 - Ao tratamento de dados pessoais efetuado através dos sistemas previstos no presente capítulo é aplicável a Lei n.º 59/2019, de 8 de agosto, e, quanto aos dados constantes do processo penal, a legislação processual penal, com as especificidades constantes dos números seguintes.
2 - Os dados biométricos relativamente aos quais o sistema não indique correspondência com as pessoas visadas são eliminados de imediato e de forma automática, não podendo ser objeto de qualquer tratamento posterior à comparação.
3 - Os dados biométricos indicados pelo sistema como correspondência que não seja confirmada nos termos dos n.ºs 4 e 5 do artigo 31.º são conservados de forma pseudonimizada e com acesso restrito, pelo prazo máximo de seis meses, exclusivamente para controlo da exatidão e deteção de enviesamentos do sistema e para o exercício do direito de defesa. Findo esse prazo, os dados são eliminados.
4 - Sem prejuízo do disposto nos n.ºs 4 e 5 do artigo 26.º e do n.º 5 do artigo 29.º, os dados e resultados que não sejam integrados em processo penal são eliminados no prazo máximo de 90 dias após a cessação da utilização.
5 - Os registos cronológicos previstos no artigo 27.º da Lei n.º 59/2019, de 8 de agosto, são conservados pelo prazo de três anos, podendo ser conservados por período mais longo quando sejam necessários a procedimentos de controlo já iniciados.
6 - Em tudo o que não estiver previsto no presente capítulo, é subsidiariamente aplicável o Código de Processo Penal.

CAPÍTULO V
Regime sancionatório

Artigo 34.º
Disposições gerais

1 - O presente capítulo estabelece o regime sancionatório aplicável às infrações ao RIA cometidas pelos operadores e pelos organismos notificados, em execução do artigo 99.º do mesmo Regulamento.
2 - As contraordenações previstas no presente capítulo classificam-se, em função da gravidade da infração, em muito graves, graves e leves.
3 - As coimas previstas no presente capítulo têm como limite máximo os montantes e as percentagens fixados nos n.ºs 3 a 5 do artigo 99.º do RIA, não sendo fixados limites mínimos especiais.
4 - Na determinação da medida concreta da coima, a autoridade competente tem em conta todas as circunstâncias relevantes do caso e, designadamente, os critérios previstos no n.º 7 do artigo 99.º do RIA.
5 - A negligência é punível, não sendo aplicável a redução do limite máximo da coima prevista no artigo 17.º do regime geral do ilícito de mera ordenação social (RGCO), aprovado pelo Decreto-Lei n.º 433/82, de 27 de outubro, sem prejuízo da ponderação do caráter negligente da infração nos termos da alínea i) do n.º 7 do artigo 99.º do RIA.
6 - Sempre que a contraordenação resulte da omissão de um dever, a aplicação da sanção e o pagamento da coima não dispensam o infrator do seu cumprimento, se este ainda for possível.
7 - O disposto no presente capítulo não prejudica:

a) A competência exclusiva da Comissão Europeia para aplicar coimas aos prestadores de modelos de IA de finalidade geral, nos termos do artigo 101.º do RIA;
b) A responsabilidade civil, penal ou disciplinar a que haja lugar.

Artigo 35.º
Contraordenações muito graves

1 - Constitui contraordenação muito grave a violação das seguintes proibições de práticas de IA previstas no n.º 1 do artigo 5.º do RIA:

a) A colocação no mercado, a colocação em serviço e a utilização de sistemas de IA que recorram a técnicas subliminares, manipuladoras ou enganadoras, nos termos da alínea a);
b) A exploração de vulnerabilidades de pessoas ou grupos de pessoas, nos termos da alínea b);
c) A geração ou manipulação de conteúdos íntimos ou sexualmente explícitos de pessoa identificável sem o seu consentimento, nos termos da alínea b-A);
d) A geração ou manipulação de material de abuso sexual de criança, nos termos da alínea b-B);
e) A avaliação ou classificação social de pessoas singulares, nos termos da alínea c);

f) A avaliação do risco de prática de infração penal com base exclusivamente na definição de perfis ou em traços de personalidade, nos termos da alínea d);
g) A criação ou expansão de bases de dados de reconhecimento facial através da recolha aleatória de imagens faciais, nos termos da alínea e);
h) A inferência de emoções no local de trabalho ou em instituições de ensino, nos termos da alínea f);
i) A categorização biométrica para deduzir ou inferir categorias sensíveis de dados, nos termos da alínea g);
j) A utilização de sistemas de identificação biométrica à distância "em tempo real" em espaços acessíveis ao público para efeitos de aplicação da lei, sem autorização, fora dos termos da autorização concedida ou em violação do disposto nos artigos 23.º a 28.º.

2 - A contraordenação muito grave é punível com coima até 35 000 000 EUR ou, tratando-se de empresa, até 7% do respetivo volume de negócios anual a nível mundial no exercício anterior, consoante o que for mais elevado.

Artigo 36.º
Contraordenações graves

1 - Constitui contraordenação grave o incumprimento, pelos prestadores de sistemas de IA de risco elevado, das seguintes obrigações previstas no RIA:

a) Observância dos requisitos aplicáveis aos sistemas de IA de risco elevado, previstos na alínea a) do artigo 16.º e nos artigos 8.º a 15.º;
b) Indicação do nome, designação comercial registada ou marca registada e endereço de contacto do prestador, prevista na alínea b) do artigo 16.º;
c) Existência de sistema de gestão da qualidade, prevista na alínea c) do artigo 16.º e no artigo 17.º;
d) Conservação da documentação, prevista na alínea d) do artigo 16.º e no artigo 18.º;
e) Conservação dos registos gerados automaticamente, prevista na alínea e) do artigo 16.º e no artigo 19.º;
f) Sujeição ao procedimento de avaliação da conformidade, prevista na alínea f) do artigo 16.º e no artigo 43.º;
g) Elaboração da declaração UE de conformidade, prevista na alínea g) do artigo 16.º e no artigo 47.º;
h) Aposição da marcação CE, prevista na alínea h) do artigo 16.º e no artigo 48.º;
i) Registo na base de dados da UE, previsto na alínea i) do artigo 16.º e no n.º 1 do artigo 49.º;
j) Adoção de medidas corretivas e prestação de informações, prevista na alínea j) do artigo 16.º e no artigo 20.º;
k) Demonstração da conformidade a pedido de autoridade nacional competente e cooperação com esta, prevista na alínea k) do artigo 16.º e no artigo 21.º;
l) Cumprimento dos requisitos de acessibilidade, previsto na alínea l) do artigo 16.º;

m) Nomeação de mandatário, tratando-se de prestador estabelecido em país terceiro, prevista no n.º 1 do artigo 22.º;
n) Acompanhamento pós-comercialização, previsto no artigo 72.º;
o) Comunicação de incidentes graves, prevista no artigo 73.º;
p) Classificação correta do sistema de IA quanto ao risco, nos termos do n.º 7 do artigo 80.º.

2 - Constitui contraordenação grave o incumprimento, pelos mandatários, das seguintes obrigações previstas no artigo 22.º do RIA:

a) Verificação da declaração UE de conformidade, da documentação técnica e da realização do procedimento de avaliação da conformidade, prevista na alínea a) do n.º 3;
b) Conservação da documentação à disposição das autoridades, prevista na alínea b) do n.º 3;
c) Prestação de informações e documentação a pedido das autoridades, prevista na alínea c) do n.º 3;
d) Cooperação com as autoridades, prevista na alínea d) do n.º 3;
e) Registo ou verificação das informações registadas, previsto na alínea e) do n.º 3;
f) Cessação do mandato e informação às autoridades, prevista no n.º 4.

3 - Constitui contraordenação grave o incumprimento, pelos importadores, das seguintes obrigações previstas no artigo 23.º do RIA:

a) Verificações prévias à colocação no mercado, previstas no n.º 1;
b) Não colocação no mercado de sistemas não conformes e informação em caso de risco, prevista no n.º 2;
c) Indicação do nome, designação comercial registada ou marca registada e endereço de contacto do importador, prevista no n.º 3;
d) Condições de armazenamento e de transporte, previstas no n.º 4;
e) Conservação da documentação, prevista no n.º 5;
f) Prestação de informações e documentação a pedido das autoridades, prevista no n.º 6;
g) Cooperação com as autoridades, prevista no n.º 7.

4 - Constitui contraordenação grave o incumprimento, pelos distribuidores, das seguintes obrigações previstas no artigo 24.º do RIA:

a) Verificações prévias à disponibilização no mercado, previstas no n.º 1;
b) Não disponibilização no mercado de sistemas não conformes e informação em caso de risco, prevista no n.º 2;
c) Condições de armazenamento e de transporte, previstas no n.º 3;
d) Adoção de medidas corretivas e informação às autoridades, prevista no n.º 4;
e) Prestação de informações e documentação a pedido das autoridades, prevista no n.º 5;
f) Cooperação com as autoridades, prevista no n.º 6.

5 - Constitui contraordenação grave o incumprimento, pelos responsáveis pela implantação de sistemas de IA de risco elevado, das seguintes obrigações previstas no RIA:

a) Utilização do sistema de acordo com as instruções de utilização, prevista no n.º 1 do artigo 26.º;
b) Atribuição da supervisão humana a pessoas com a competência, a formação e a autoridade necessárias, prevista no n.º 2 do artigo 26.º;
c) Pertinência e representatividade dos dados de entrada, prevista no n.º 4 do artigo 26.º;
d) Controlo do funcionamento do sistema, informação e suspensão da utilização em caso de risco ou de incidente grave, previstos no n.º 5 do artigo 26.º;
e) Conservação dos registos gerados automaticamente, prevista no n.º 6 do artigo 26.º;
f) Informação aos representantes dos trabalhadores e aos trabalhadores afetados, prevista no n.º 7 do artigo 26.º;
g) Registo, tratando-se de autoridade pública, previsto no n.º 8 do artigo 26.º;
h) Utilização das informações do prestador na avaliação de impacto sobre a proteção de dados, prevista no n.º 9 do artigo 26.º;
i) Autorização e condições de utilização de sistemas de identificação biométrica à distância em diferido, previstas no n.º 10 do artigo 26.º do RIA e nos artigos 29.º e 30.º da presente lei;
j) Informação às pessoas singulares sujeitas à utilização do sistema, prevista no n.º 11 do artigo 26.º;
k) Cooperação com as autoridades, prevista no n.º 12 do artigo 26.º;
l) Avaliação de impacto sobre os direitos fundamentais e notificação dos resultados, previstas no artigo 27.º;
m) Prestação de explicações às pessoas afetadas, prevista no artigo 86.º.

6 - Constitui contraordenação grave o incumprimento, pelos organismos notificados, das seguintes obrigações previstas no RIA:

a) Requisitos aplicáveis aos organismos notificados, previstos no artigo 31.º;
b) Subcontratação e recurso a filiais, previstos nos n.ºs 1, 3 e 4 do artigo 33.º;
c) Obrigações operacionais, previstas no artigo 34.º.

7 - Constitui contraordenação grave o incumprimento, pelos prestadores e pelos responsáveis pela implantação, das seguintes obrigações de transparência previstas no artigo 50.º do RIA:

a) Informação sobre a interação com um sistema de IA, prevista no n.º 1;
b) Marcação de conteúdos sintéticos de áudio, imagem, vídeo ou texto, prevista no n.º 2;
c) Informação sobre a utilização de sistemas de reconhecimento de emoções ou de categorização biométrica, prevista no n.º 3;
d) Divulgação de falsificações profundas e de texto artificialmente gerado ou manipulado sobre questões de interesse público, prevista no n.º 4;
e) Forma e momento da prestação das informações, previstos no n.º 5.

8 - A contraordenação grave é punível com coima até 15 000 000 EUR ou, tratando-se de empresa, até 3% do respetivo volume de negócios anual a nível mundial no exercício anterior, consoante o que for mais elevado.

Artigo 37.º
Contraordenações leves

1 - Constitui contraordenação leve a prestação de informações inexatas, incompletas ou enganosas aos organismos notificados ou às autoridades nacionais competentes, designadamente:

a) Em resposta a pedido de informação ou de documentação formulado por uma autoridade de fiscalização do mercado ou pela autoridade notificadora;
b) No âmbito de procedimento de avaliação da conformidade conduzido por um organismo notificado.

2 - A contraordenação leve é punível com coima até 7 500 000 EUR ou, tratando-se de empresa, até 1% do respetivo volume de negócios anual a nível mundial no exercício anterior, consoante o que for mais elevado.

Artigo 38.º
Regime especial das micro, pequenas e médias empresas

1 - Tratando-se de micro, pequena ou média empresa, incluindo as empresas em fase de arranque, ou de pequena empresa de média capitalização, aplica-se, para cada uma das contraordenações previstas nos artigos 35.º a 37.º, o limite mais baixo entre o montante fixo e a percentagem do volume de negócios aí indicados.

2 - Para efeitos do presente artigo, as micro, pequenas e médias empresas são as definidas na Recomendação 2003/361/CE da Comissão, de 6 de maio de 2003, e as pequenas empresas de média capitalização são as definidas no anexo da Recomendação (UE) 2025/1099 da Comissão, de 21 de maio de 2025.

Artigo 39.º
Autoridades e organismos públicos

1 - Às autoridades e organismos públicos que cometam as infrações previstas nos artigos 35.º a 37.º é aplicada admoestação, acompanhada da determinação das medidas necessárias à cessação da infração e do respetivo prazo de cumprimento, não lhes sendo aplicáveis coimas.

2 - O disposto no número anterior não prejudica a aplicação das sanções acessórias previstas nas alíneas a) a c) do n.º 1 do artigo seguinte, nem o exercício dos poderes de fiscalização previstos no RIA.

3 - A autoridade competente comunica a verificação da infração ao membro do Governo responsável pela tutela ou superintendência da entidade infratora, para efeitos de apuramento de eventual responsabilidade disciplinar.

4 - O disposto no presente artigo não se aplica às entidades públicas empresariais, que ficam sujeitas ao regime aplicável às empresas.

Artigo 40.º
Sanções acessórias

1 - Consoante a gravidade da infração e a culpa do agente, podem ser aplicadas, cumulativamente com a coima ou com a admoestação prevista no artigo 39.º, as seguintes sanções acessórias:

a) Apreensão do sistema de IA ou determinação da sua retirada do mercado;
b) Suspensão da colocação em serviço do sistema até sanação das não conformidades identificadas;
c) Publicidade da decisão condenatória, a expensas do infrator;
d) Interdição temporária do exercício da atividade, em caso de reincidência em contraordenação muito grave.

2 - A sanção acessória prevista na alínea d) do número anterior tem a duração máxima de dois anos, contados a partir da decisão condenatória definitiva.

3 - A aplicação de sanções acessórias respeita os princípios da proporcionalidade e da adequação aos fins de prevenção geral e especial prosseguidos.

Artigo 41.º
Competência sancionatória

1 - São competentes para a instauração, instrução e decisão dos processos de contraordenação previstos na presente lei:

a) A autoridade de fiscalização do mercado competente, designada nos termos do artigo 9.º;
b) A autoridade notificadora, quanto às infrações previstas no n.º 6 do artigo 36.º.

2 - A instauração e a decisão dos processos de contraordenação, incluindo a aplicação de coimas e sanções acessórias, são da competência do órgão de administração da autoridade competente ou, na sua falta, do respetivo dirigente máximo.

3 - A instrução dos processos de contraordenação cabe aos serviços da autoridade competente.

4 - As competências previstas no n.º 2 podem ser delegadas, com possibilidade de subdelegação, nos termos dos estatutos ou da lei orgânica da autoridade competente.

5 - Quando a mesma conduta seja suscetível de ser apreciada por mais do que uma das autoridades referidas no n.º 1, aplica-se o disposto no artigo 14.º.

Artigo 42.º
Concurso de infrações e proibição de dupla sanção

1 - Sempre que a mesma conduta constitua simultaneamente infração ao RIA e a outro regime sancionatório, designadamente o regime de proteção de dados pessoais, o regime dos serviços digitais ou os regimes sancionatórios da supervisão financeira, o agente não pode ser sancionado duas vezes pelos mesmos factos e pelo mesmo fundamento.

2 - Nos casos previstos no número anterior, as autoridades competentes ao abrigo de cada um dos regimes determinam, de comum acordo, qual delas instrui o processo, tendo em conta

a natureza predominante da infração, aplicando-se, na falta de acordo, o disposto no artigo 14.º, com as necessárias adaptações.

Artigo 43.º
Impugnação judicial

1 - A impugnação judicial das decisões, despachos e demais medidas adotados pelas autoridades competentes em processo de contraordenação previsto na presente lei, ou em diligências que o precedam, rege-se pelos artigos 55.º e 59.º e seguintes do RGCO, sendo competente o Tribunal da Concorrência, Regulação e Supervisão.
2 - Das decisões do Tribunal da Concorrência, Regulação e Supervisão cabe recurso para o Tribunal da Relação de Lisboa, que decide em última instância.
3 - A autoridade que proferiu a decisão participa no processo de impugnação na qualidade de sujeito processual e goza dos respetivos direitos, incluindo na audiência de julgamento.
4 - O disposto nos números anteriores aplica-se independentemente da autoridade competente, ainda que esta disponha de regime próprio de impugnação noutros domínios.

Artigo 44.º
Afetação do produto das coimas

O produto das coimas aplicadas nos termos da presente lei reverte:

a) Em 60% para o Estado;
b) Em 40% para a autoridade que aplicou a coima, destinando-se este montante ao reforço da respetiva capacidade técnica de supervisão dos sistemas de IA.

Artigo 45.º
Direito subsidiário

Em tudo o que não estiver especialmente previsto no presente capítulo, é subsidiariamente aplicável, com as necessária168/2013 adaptações, o RGCO, na sua redação atual.

CAPÍTULO VI
Disposições transitórias e finais

Artigo 46.º
Proteção de dados pessoais

1 - A presente lei não prejudica o Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016 (RGPD), as Leis n.ºs 58/2019 e 59/2019, ambas de 8 de agosto, e a Lei n.º 41/2004, de 18 de agosto, nem as atribuições da CNPD.
2 - As autoridades nacionais competentes e a CNPD cooperam entre si e trocam as informações necessárias ao exercício das respetivas atribuições.

Artigo 47.º

Categorias especiais de dados para correção de enviesamentos

1 - Para efeitos do artigo 6.º da Lei n.º 59/2019, de 8 de agosto, considera-se autorizado por lei o tratamento de categorias especiais de dados pessoais pelas autoridades competentes, na aceção dessa lei, quando efetuado nos termos do artigo 4.º-A do RIA e com observância das condições e garantias nele previstas.

2 - Compete à CNPD fiscalizar o cumprimento das condições e garantias previstas no artigo 4.º-A do RIA quanto ao tratamento de dados pessoais, sem prejuízo das competências das autoridades de fiscalização do mercado quanto aos requisitos do artigo 10.º do RIA.

3 - O registo e a fundamentação previstos na alínea f) do n.º 1 do artigo 4.º-A do RIA são facultados, a pedido, à CNPD e, tratando-se de sistemas de IA de risco elevado, também à autoridade de fiscalização do mercado competente.

Artigo 48.º

Norma transitória

Até à publicação da resolução do Conselho de Ministros prevista no n.º 1 do artigo 11.º, são autoridades responsáveis pela proteção dos direitos fundamentais as autoridades e os organismos públicos constantes da lista publicada pela ANACOM, nos termos do n.º 2 do artigo 77.º do RIA, à data de entrada em vigor da presente lei.

Artigo 49.º

Designação das autoridades

O Governo aprova, no prazo de 60 dias a contar da data de entrada em vigor da presente lei, o decreto-lei previsto nos artigos 8.º e 9.º e a resolução do Conselho de Ministros prevista no n.º 1 do artigo 11.º.

Artigo 50.º

Regulamentação

1 - A ANACOM aprova o regulamento previsto no n.º 8 do artigo 16.º no prazo de 90 dias a contar da data de entrada em vigor do decreto-lei referido no artigo anterior.

2 - A ANACOM aprova o regulamento previsto no n.º 4 do artigo 20.º no prazo de 90 dias a contar da data de entrada em vigor da presente lei.

Artigo 51.º

Produção de efeitos

1 - As disposições da presente lei que se reportem a obrigações previstas no RIA produzem efeitos a partir da data em que essas obrigações se tornem aplicáveis.

2 - A obrigação de inscrição prevista no n.º 2 do artigo 16.º produz efeitos a partir da data de aplicação das normas do RIA relativas aos sistemas de risco elevado mencionados no seu anexo III.

3 - O artigo 47.º produz efeitos a partir da data de aplicação do artigo 4.º-A do RIA, independentemente da data de aplicação das obrigações relativas aos sistemas de IA de risco elevado.

Artigo 52.º
Entrada em vigor

A presente lei entra em vigor no dia seguinte ao da sua publicação.



DOS BENS INFORMÁTICOS FUNDAMENTAIS À HARMONIZAÇÃO DAS POLÍTICAS DE PROMOÇÃO DA CIBERSEGURANÇA E PREVENÇÃO DO CIBERCRIME

PEDRO DIAS VENÂNCIO*

* Professor Auxiliar na Escola de Direito da Universidade do Minho. Diretor do Mestrado em Direito e Informática. Investigador integrado no JusGov (Centro de Investigação em Justiça e Governação) pdvenancio@direito.uminho.pt

RESUMO:

A nossa confiança na Sociedade da informação assenta na confidencialidade, integridade, autenticidade e disponibilidades de dados, sistemas e comunicações informáticas, como elementos estruturais essenciais à preservação da sua funcionalidade enquanto comunidade social, cultural, económica e política. O que determina que estes bens informáticos assumam um valor jurídico fundamental que se reflete, desde logo, na sua tutela constitucional.

Neste texto partimos desta constatação para identificar uma identidade de bens e valores protegidos pelas políticas de cibersegurança (na perspetiva da prevenção de vulnerabilidades e riscos) e pelas políticas de combate à cibercriminalidade (na perspetiva de repressão de condutas ilícitas). Desta identidade de valores, pensamos decorrer igualmente a necessidade de uma interdependência entre as medidas de cibersegurança de identificação e prevenção de vulnerabilidades e os atos reprimidos pelos crimes de acesso ilegítimo a sistema informático ou interceção ilegítima de comunicação informática que exploram essas mesmas vulnerabilidades.

É inegável que os agentes (*hackers*) que praticam estes atos de acesso não autorizado a sistemas e/ou comunicações eletrónicas encontram-se na posse de informação crítica sobre as vulnerabilidades de sistemas informáticos e, logo, relevante para as políticas de cibersegurança. Dependendo o sucesso das medidas de cibersegurança do acesso tempestivo desta informação crítica, defende-se ser útil encontrar mecanismos legais que promovam a partilha com as autoridades competentes e titulares de sistemas de informação relativa a vulnerabilidades dos sistemas. Inclusivamente por parte destes agentes (hackers) que desafiam a segurança de sistemas informáticos, atuando sem autorização (e, logo, ilegalmente), mas ainda assim sem intenção maliciosa ou efeitos danosos.

Tomamos neste texto como exemplo desta política legislativa a recente introdução à Lei do Cibercrime portuguesa de um novo artigo 8.º-A sob a epígrafe «*Atos não puníveis por interesse público de cibersegurança*».

Palavras-chave: Cibersegurança; cibercrime; sistemas informáticos; dados informáticos; comunicações informáticas.

ABSTRACT:

Our trust in the Information Society is based on the confidentiality, integrity, authenticity, and availability of data, systems, and computer communications, as structural elements essential to preserving its functionality as a social, cultural, economic, and political community. This means that these digital assets hold a fundamental legal value, which is reflected, first and foremost, in their constitutional protection.

In this text, I use this observation as a starting point to identify the assets and values protected by cybersecurity policies (from the perspective of preventing vulnerabilities and risks) and by policies to combat cybercrime (from the perspective of suppressing unlawful conduct).

I believe that this shared set of values also gives rise to the need for interdependence between cybersecurity measures aimed at identifying and preventing vulnerabilities and the acts criminalized under the offenses of unlawful access to a computer system or unlawful interception of computer communications that exploit those very same vulnerabilities.

It is undeniable that actors (hackers) who commit these acts of unauthorized access to systems and/or electronic communications are in possession of critical information regarding the vulnerabilities of computer systems. Information that is therefore relevant to cybersecurity policies. Since the success of cybersecurity measures depends on timely access to this critical information, it is argued that it would be useful to establish legal mechanisms that promote the sharing of information regarding system vulnerabilities with the competent authorities and information system operators. This includes information provided by those actors (hackers) who challenge the security of computer systems by acting without authorization (and, therefore, illegally), yet without malicious intent or harmful effects.

In this text, it is cited as an example of this legislative policy the recent introduction of a new Article 8-A into the Portuguese Cybercrime Law under the heading "Acts Not Punishable in the Public Interest of Cybersecurity.[1]

Keywords: Cybersecurity; cybercrime; computer systems; computer data; computer communications.

1 Tradução realizada com a ajuda da aplicação Deepl.com (versão gratuita).

I. UMA NOÇÃO DE *CIBERSEGURANÇA*

A sociedade moderna transitou para uma vivência cada vez mais digital, impulsionada pela generalização do acesso e utilização de Tecnologias da Informação e da Comunicação (TIC), em particular desde a expansão planetária da Internet[2]. Neste contexto, as crescentes capacidades de digitalização, processamento e armazenamento das TIC determinaram a transição da atividade económica, cultural e social para estes neófitos ambientes digitais, impulsionando a Era da designada Sociedade da Informação[3] enquanto novo modelo de comunidade.

Esta realidade é já presente e irreversível. Nela a segurança dos dados, dos sistemas e das comunicações informáticas assume-se como um desígnio de toda a sociedade, indispensável à preservação da sua funcionalidade enquanto comunidade social, económica e política.

Esta progressiva transição digital determinou que os Estados começassem a reconhecer nos bens digitais - dados, sistemas e comunicações informáticas - valores essenciais à sociedade. É o caso do artigo 35.º da Constituição da República Portuguesa que consagra um leque de direitos fundamentais no âmbito da utilização da informática, como sejam os dados pessoais e o direito de livre acesso às redes informáticas de uso público[4].

A este novo eixo das políticas públicas e privadas designamos de **Cibersegurança**, que seguindo uma noção dada pela *Telecomunication Standardization Sector - The International Telecommunication Union*, podemos definir como:

> «(…) *o conjunto de ferramentas, políticas, conceitos de segurança, medidas de proteção, diretrizes, abordagens de gestão de riscos, ações, formação, melhores práticas, garantias e tecnologias que podem ser utilizadas para proteger o ambiente cibernético e os ativos da organização e dos utilizadores. Os ativos da organização e dos utilizadores incluem dispositivos informáticos ligados à rede, pessoal, infraestruturas, aplicações, serviços, sistemas de telecomunicações e a*

2 Muito resumidamente "*A Internet é um sistema global de redes de computadores interligadas que utilizam um conjunto próprio de protocolos (Internet Protocol Suite ou TCP/IP) com o propósito de servir progressivamente usuários no mundo inteiro*.", in https://pt.wikipedia.org/wiki/Internet, acedido a 2-2-2022.

3 Sobre o conceito de Sociedade da Informação cfr. VENÂNCIO, Pedro Dias, A Tutela Jurídica do Formato de Ficheiro Eletrónico, Coimbra: Almedina, 2016, p. 34-36..

4 Nesse sentido, ler VENÂNCIO, Pedro Dias, "*A interoperabilidade como garantia do direito fundamental à liberdade de acesso às redes informatizadas de uso público, previsto no artigo 35.o n.o 6 da C.R.P.*", Anuário de Direitos Humanos, v. 2, p. 71-88, 2019.

totalidade da informação transmitida e/ou armazenada no ambiente cibernético. A cibersegurança visa garantir a obtenção e a manutenção das propriedades de segurança dos ativos da organização e dos utilizadores face aos riscos de segurança relevantes no ambiente cibernético.

Os objetivos gerais de segurança incluem o seguinte: Disponibilidade; Integridade, que pode incluir autenticidade e não repúdio; e Confidencialidade.»[5]

Em conclusão, as políticas de cibersegurança definem os procedimentos que visam proteger t*rês bens jurídicos essenciais* (dados, sistemas e comunicações informáticas), para os quais se pretende garantir *quatro valores* (confidencialidade, integridade, autenticidade e disponibilidades), que exporemos sucintamente nas secções seguintes.

II. OS BENS JURÍDICOS NUCLEARES DA SOCIEDADE DA INFORMAÇÃO

Dissemos supra que a cibersegurança visa proteger três bens jurídicos nucleares da Sociedade da Informação e que são (1) os *sistemas informáticos*, os (2) *dados informáticos*, e as (3) *comunicações informáticas*. Apresentaremos nesta secção uma curta noção destes bens jurídicos.

O conceito de **sistemas informático** é atualmente definido na legislação portuguesa pela Lei do Cibercrime[6], no seu artigo 1.º alínea a), como:

«*qualquer dispositivo ou conjunto de dispositivos interligados ou associados, em que um ou mais de entre eles desenvolve, em execução de um programa, o tratamento automatizado de dados informáticos, bem como a rede que suporta a comunicação entre eles e o conjunto de dados informáticos armazenados, tratados, recuperados ou transmitidos por aquele ou aqueles dispositivos, tendo em vista o seu funcionamento, utilização, proteção e manutenção*;»

5 Recommendation ITU-T X.1205 (04/2008), da Telecomunication Standardization Sector - The International Telecommunication Union, disponível em https://www.itu.int/rec/t-rec-x.1205-200804-i (acedido em 27/06/2026). (Tradução da nossa responsabilidade com a ajuda da versão gratuita do tradutor - DeepL.com).

6 Aprovada pela Lei n.º 109/2009, de 15 de setembro

O conceito de sistema informático é de um objeto complexo[7], na medida em que incluí não só todo o *hardware* e *software* que estando interligados cumprem as funções para que estão desenhados, mas também a rede de comunicações que suportam essas interligações e os dados nele partilhados. O sistema informático integra o complexo e multifacetado conjunto das realidades corpóreas e incorpóreas que o tornam um *mundo digital* funcional.

É também um conceito neutro, pois para os fins da Lei do Cibercrime (como para a cibersegurança) é irrelevante a sua dimensão, a pessoas ou entidade a que pertence, ou a relevância social, económica ou cultural das funções que desempenha. Todo o sistema informático é protegido.

Também na Lei do Cibercrime portuguesa [artigo 1.º alínea d)] encontramos a definição que nos ajuda a compreender o conceito de **dado informático**:

> «*qualquer representação de factos, informações ou conceitos sob uma forma suscetível de processamento num sistema informático, incluindo os programas aptos a fazerem um sistema informático executar uma função*»

Desta definição podemos depreender que o conceito de dado informático abarca qualquer *registo em formato eletrónico*[8] independentemente do seu conteúdo inteligível, por humanos ou pelo próprio sistema informático, incluindo os programas informáticos[9].

Para os fins gerais da cibersegurança, o conceito de dado informático abstrai-se da natureza da informação nele registada e assume-se como um bem jurídico que reclama uma tutela autónoma, face aquela que possa incidir sobre a informação nele representada. A cibersegurança protege o próprio direito a registar e conservar dados em formato informático como um bem jurídico essencial à mundividência digital das sociedades modernas.

Por fim, o conceito de **comunicações informáticas** encontra o seu paralelo no conceito de rede informática que assenta na possibilidade de diferentes aparelhos informáticos poderem trocar entre si dados informáticos, independentemente do seu conteúdo informativo, através destas tecnologias de comunicação comuns. Estas tecnologias que permitem esta transmissão

---

7 Damos uma noção aprofundada do conceito de Sistema Informático em VENÂNCIO, Pedro Dias, Manual de Propriedade Intelectual Digital, 1.a. Coimbra: Editora D'Ideias, 2023, p. 91 e ss.

8 Em termos técnicos este registo em formato eletrónico é usualmente designada por "ficheiro eletrónico" que abordarmos de forma mais extensa em *Ibid*., p. 303-311.

9 Assim, um dado informático pode representar um dado pessoal (para os fins da legislação de proteção de dados pessoais); uma obra ou prestação protegidas por direito de autor ou direitos conexos; uma invenção protegida por um registo de patente; um sinal protegido por um registo de marca; um segredo de negócios; um contrato; ou até um valor patrimonial (seja em moeda de curso legal, seja um ativo digital transacionável), etc.

de dados dentro e entre sistemas informáticos são os designados protocolos de comunicação[10]. Podemos definir um protocolo de comunicação como um sistema de regras de sintaxe, semântica e sincronização que permitem a comunicação entre computadores[11].
Quando nos referimos a comunicações informáticas como um bem jurídico protegido pelas medidas de cibersegurança, estamos a referir-nos não só à preservação e segurança da funcionalidade técnica dos respetivos protocolos de comunicação (ao processo de transmissão eletrónica de dados propriamente dito), mas também à preservação da integridade e confidencialidade dos conteúdos informativos transmitidos (independentemente da sua natureza).

III. OS VALORES A CONSERVAR NOS BENS JURÍDICOS PROTEGIDOS

Da parte final da definição que supra deixamos de cibersegurança transparecem os valores essenciais que esta visa assegurar, usualmente aceites pela doutrina e designados pela sigla *CIA*, da trilogia anglo-saxónica *Confidentiality, Integrity, & Availability* (confidencialidade, integridade e disponibilidade).
Nós tendemos a acrescentar um quarto valor, que pensamos ser possível destacar da *integridade* como um valor autónomo, e que designamos de *autenticidade*.
Explicaremos melhor a seguir estes valores, bem como porque entendemos pertinente destacar a autenticidade como um valor autónomo.
A **confidencialidade**[12] representa a manutenção de dados, sistemas e comunicações conhecidos apenas dos sujeitos com legitimidade de acesso. A confidencialidade protege

10 Alguns dicionários definem protocolo com apenas uma «*norma regulando a transmissão de dados entre computadores*» [DOWNING, Douglas *et al*, Dicionário de Termos Informáticos e da Internet, Lisboa: Paralelo Editora, Lda, 2001, p. 470.], ou o «*conjunto de regras que descrevem a forma como é efetuada a transmissão e receção de dados, nomeadamente numa rede informática*» [MATOS, José António Alves de, Dicionário de Informática e Novas Tecnologias, LISBOA: Lidel - Edições Técnicas, L.da, 2009, p. 296.].
11 Nesse sentido refere José Alberto Vieira que «os protocolos de acesso podem ser definidos como a "*série de regras sintáticas e semânticas que determinam o comportamento de unidades funcionais na obtenção de comunicação*" [VIEIRA, José Alberto, A protecção dos programas de computador pelo Direito de Autor, [s.l.]: Lex, 2005, p. 440.]. Aprofundamos este conceito numa perspetiva técnica em VENÂNCIO, Manual de Propriedade Intelectual Digital, p. 375-377.
12 Embora também designada por privacidade por alguns autores, nos preferimos a expressão confidencialidade, por a entendermos com um conceito mais abrangente. A privacidade está usualmente reservada à tutela das pessoas, dos seus dados pessoais, da sua intimidade, etc… A confidencialidade pode reportar-se aos dados pessoais e íntimos de pessoas singulares, mas também aos dados comerciais de empresas,

contra a intromissão de terceiros (acesso não autorizado) em dados, sistemas ou comunicações informáticas.

A **integridade** representa a exatidão, consistência e fiabilidade dos dados ao longo de todo o seu ciclo de vida. Isto significa que o sistema informático deve estar desenhado de modo a impedir que dados, comunicações ou o próprio sistema possam ser inutilizados, danificados ou indevidamente alterados, seja por entidades não autorizadas, seja por eventos fortuitos ou maliciosos. O que significa que a integridade se relaciona não só com o acesso e manipulação não autorizada de dados, comunicações e sistemas, mas também com a sua deterioração por erros ou eventos fortuitos.

É por força desta definição de integridade, usual na doutrina e ciência informática, que entendemos que se deve autonomizar a autenticidade. Para nós, a **autenticidade** significa - enquanto valor autónomo para a cibersegurança - não apenas o dano ou adulteração de dados genuínos (que se poderia caracterizar apenas como uma perda de integridade), mas também que os dados não foram criados por quem não tinha legitimidade para o fazer - o que contende com a verificação e autentificação da entidade dos sujeitos atuantes em ambiente digital - e/ou que não representam factos ou informações falsas, independentemente de quem e como foram criados - o que se relaciona com a confiança jurídica nos dados, sistemas e comunicações eletrónicas.

Por fim, a **disponibilidade** representa a possibilidade dos legítimos utilizadores terem acesso em tempo útil e sem condicionamentos ou limitações aos seus dados, sistemas e comunicações eletrónicas.

IV. AS PRINCIPAIS MEDIDAS DE CIBERSEGURANÇA

Embora a implementação efetiva de uma política de cibersegurança possa variar em função da dimensão da organização ou comunidade onde se pretende implementar, há um núcleo essencial de medidas que constituem a base das boas práticas de cibersegurança.

ou a quaisquer outros dados, de pessoas ou entidades coletivas, que estes quem manter reservados por qualquer motivo legítimo.

Os padrões internacionalmente reconhecidos, adotados por exemplo no Quadro Nacional de Referência para a Cibersegurança[13], apontam para cinco medidas de cibersegurança essenciais[14] - *identificar, proteger, detetar, responder e recuperar* - em volta das quais se irão construir as estratégias, políticas e regimes legais de cibersegurança.

Seguindo de perto as definições dadas no citado Quadro Nacional de Referência para a Cibersegurança[15] podemos definir estas cinco medidas da seguinte forma:

1. **Identificar** as vulnerabilidades e riscos de cada sistema informático, o que significa compreender «*o contexto da organização, dos ativos que suportam os processos críticos da atividade da organização e dos riscos associados relevantes. Esta compreensão permite que a organização consiga definir e priorizar os seus recursos e investimentos, de acordo com os seus objetivos gerais e com a sua estratégia de gestão do risco*»;
2. **Proteger** ou prevenir a ocorrência dos riscos identificados, significa a «*implementação de medidas destinadas a proteger os processos organizativos e os ativos da organização, independentemente da sua natureza tecnológica. Assim, nesta categoria, são definidas medidas orientadas à proteção da organização nas suas três dimensões: Pessoas, Processos e Tecnologia*»;
3. **Detetar** significa a «*definição e implementação de medidas destinadas a identificar, de forma atempada, os incidentes. Ou seja, a deteção de eventos com um efeito adverso real na segurança das redes e dos sistemas de informação*»;
4. **Responder** significa a «*definição e implementação de medidas de ação apropriadas, em caso de deteção de um incidente. As medidas propostas no âmbito deste objetivo pretendem mitigar o impacto do incidente, ou seja, reduzir os seus potenciais efeitos adversos*»; e
5. **Recuperar** significa a «*definição e implementação de atividades, que visam a gestão de planos e medidas de recuperação dos processos e serviços afetados por um incidente de cibersegurança. As medidas pertencentes a este objetivo*

13 CNCS, Centro Nacional de Cibersegurança, *Quadro Nacional de Referência para a Cibersegurança*, [s.l.]: Gabinete Nacional de Segurança, 2019, p. 47-48.

14 A estas cinco medidas, acrescenta-se modernamente uma sexta - a governança - que representa o patamar organizacional da implementação destas medidas. Não iremos abordar este patamar.

15 CNCS, *Quadro Nacional de Referência para a Cibersegurança*, p. 47-48..

pretendem assegurar a resiliência da organização nas suas dimensões: Pessoas, Processos e Tecnologia. E que, no caso de existência de um incidente, a organização consiga utilizar as medidas para suporte à recuperação em tempo útil da sua atividade».

V. A TUTELA DA BENS DIGITAIS NOS CRIMES INFORMÁTICOS EM SENTIDO ESTRITO

Temos definido a criminalidade informática em sentido estrito "*como aquela em que o elemento digital surge como parte integradora do tipo legal ou mesmo como seu objeto de proteção. Ou seja, não só os que têm por bem jurídico protegido o próprio acesso ou funcionalidade da sociedade da informação, mas todos aqueles em que a informática faz parte necessária dos seus elementos típicos*"[16].

O núcleo essencial destes crimes informáticos em sentido estrito encontra-se previsto nos tipos legais originalmente consagrados na Convenção sobre o Cibercrime, adotada em Budapeste em 23 de novembro de 2001, e dirigem-se especificamente à tutela dos bens jurídicos relevantes para a cibersegurança. Na verdade, neste conjunto nuclear de tipos legais de crimes informáticos em sentido estrito (os crimes de *acesso ilegítimo a sistema informático*; *interceção ilegítima de comunicações eletrónicas*; *dano relativo a dados informáticos*; a *falsidade de dados informáticos*; e a *sabotagem de sistema informático*[17]) encontra-se a tutela daqueles que são os bens e valores jurídicos essenciais que as medidas de cibersegurança visam acautelar: a integridade, autenticidade, confidencialidade e disponibilidade de dados, comunicações e sistemas eletrónicos[18].

16 VENÂNCIO, Pedro Dias, Lições de Direito do Cibercrime. E da tutela penal de dados pessoais, 2.a. [s.l.]: Editora D'Ideias, 2025, p. 24.

17 Para além destes cinco tipos nucleares existem outros que, direta ou indiretamente, tutelam igualmente bens jurídicos essenciais da cibersegurança. Optamos por limitar a nossa análise a estes cinco tipos legais de crimes informáticos, desde logo por serem os que estão harmonizados internacionalmente de forma mais abrangente, mas também por serem os que se ligam de forma mais umbilical aos valores da cibersegurança.

18 Como defendemos nas nossas Lições de Direito do Cibercrime: «*Numa terceira categoria, decorrente da elevação dos dados, redes e sistemas informáticos a bens e serviços essenciais à comunidade, surge um conjunto de crimes que tutelam a própria confiança, fiabilidade, e segurança de redes e sistemas informáticos e/ou a integridade dos bens digitais. Estamos aqui no âmbito do que nos parece constituir o núcleo essencial da criminalidade informática em sentido estrito*» [VENÂNCIO, Lições de Direito do Cibercrime. E da tutela penal de dados pessoais, p. 26.]

Analisaremos de seguida, sucintamente, a forma simples destes tipos de crimes para neles identificarmos os bens jurídicos que se visam salvaguardar pelas medidas de cibersegurança.

O **crime de aceso ilegítimo a sistema informático**[19], tem dois elementos objetivos no tipo legal: (1) o agente atuar não estando legitimado pelo titular do sistema informático e (2) praticar um ato de aceder a um sistema informático ou de parte dele. Os tratados internacionais não exigem um elemento subjetivo de dolo específico (como seja a "intenção de alcançar, para si ou para outrem, um benefício ou vantagem ilegítimos" presente em outros tipos legais), embora não o proíbam, ficando essa possibilidade dentro da liberdade de conformação dos Estado aderente. Não se prevendo esse elemento subjetivo específico, como é o caso da lei portuguesa, este crime torna-se extraordinariamente abrangente nas condutas puníveis[20].

Como vem salientando a jurisprudência portuguesa «*O bem jurídico protegido é a segurança do sistema informático*»[21]. Diríamos nós, seguindo a terminologia adotada nesta obra, que o crime de acesso ilegítimo tutela a **confidencialidade dos sistemas informáticos**.

O **crime de dano relativo a dados informáticos**[22], consuma-se atualmente com apenas dois elementos objetivos: (1) o agente atuar não estando legitimado; e (2) praticar um ato de apagar, alterar, destruir, no todo ou em parte, danificar, suprimir ou tornar não utilizáveis ou não acessíveis programas ou outros dados informáticos alheios ou por qualquer forma lhes afetar a capacidade de uso. Uma vez mais, os tratados internacionais não exigem que se

19 Em Portugal previsto no artigo 6.º da Lei do Cibercrime. Sobre o crime de acesso ilegítimo , entre outros: SIMÕES, Pedro Dias, O "*Hacking*" enquanto Crime de Acesso Ilegítimo. Das suas especialidades à utilização das mesmas para a fundamentação de um novo direito., Direito da Sociedade da Informação, v. VIII, p. 229 a 261, 2009; Pedro Verdelho em anotação ao art.o 6.o, VV., Comentário das Leis Penais Extravagantes - Volume 1, Lisboa: Universidade Católica Portuguesa, 2010, p. 515-517; FREITAS, Pedro, Breves nótulas sobre o crime de acesso ilegítimo previsto na Lei do Cibercrime, *in*: VV, AA (Org.), Estudos em Comemoração dos 20 anos da Escola de Direito da Escola de Direito da Universidade do Minho, Coimbra: Coimbra Editora, 2014, p. 565–585; VENÂNCIO, Lições de Direito do Cibercrime. E da tutela penal de dados pessoais, p. 73-77; NUNES, Duarte Rodrigues, Os Crimes Previstos na Lei do Cibercrime. E a responsabilidade penal dos entes colectivos**.**, 2.a. Coimbra: GestLegal, 2024.

20 Esta nova amplitude do tipo legal questão foi muito debatida entre a comunidade cibernética ligada às tecnologias da informação e comunicação, em especial os movimentos "*hacker*", que contestam a penalização dos acessos não maliciosos inseridos no âmbito dos designados "*white hat hacker*" ou "*ethical hackers*", e que em certa medida justifica a recente introdução à Lei do Cibercrime do artigo 8.º-A relativo a atos não puníveis por interesse público de cibersegurança, e que trataremos na secção seguinte.

21 Do sumário do acórdão do Tribunal da Relação do Porto de 08-01-2014, no processo n.º 1170/09.8JAPRT.P2, sendo relatora Elsa Paixão, in http://www.dgsi.pt.

22 Em Portugal previsto no artigo 4.º da LC. Sobre o crime de dano relativo a programas ou outros dados informáticos na Lei do Cibercrime ler: Pedro Verdelho em anotação ao art.o 4.o, VV., Comentário das Leis Penais Extravagantes - Volume 1, p. 509-512; VENÂNCIO, Lições de Direito do Cibercrime. E da tutela penal de dados pessoais; NUNES, Os Crimes Previstos na Lei do Cibercrime, p. 67-120.

preveja um elemento subjetivo de dolo específico, nem o legislador português o fez. Temos defendido o bem jurídico protegido pelo crime de dano informático é a **integridade dos dados informáticos** e, indiretamente, o património digital do lesado[23].

O **crime de falsidade de dados informáticos**[24] na sua forma simples vem previsto no n.º 1 artigo 3.º da LC e tem como elementos objetivos do tipo legal: (1) o agente praticar a ação de, por qualquer outra forma, interferir num tratamento informático de dados e (2) o efeito de produzir dados ou documentos não genuínos. Neste caso incluísse no tipo legal como elementos subjetivos do crime a intenção de provocar engano nas relações jurídicas, no sentido de que estes (*os dados ou documentos não genuínos*) sejam considerados ou utilizados para finalidades juridicamente relevantes como se fossem verdadeiros.

Recorrendo aos valores da Sociedade da Informação que supra, se nos ficássemos pela tríade *CIA* diríamos que o bem protegido seria a *integridade dos dados informáticos*, mas com a autonomizamos como valor próprio, diremos que a falsidade informática tutela a **autenticidade dos dados informáticos**.

A ação ilícita prevista no tipo legal **de crime de sabotagem de sistema informático**[25] é em tudo similar aos elementos objetivos do crime de dano informático nos seus dois primeiros elementos objetivos: (1) o agente atuar não estando legitimado; e (2) praticar atos de "de qualquer outra forma de interferência em sistema informático", aos quais se acrescenta, como *efeito* (ou resultado) necessário do tipo, o (3) entravar, impedir, interromper ou perturbar gravemente o funcionamento de um sistema informático.

Em coerência com a terminologia que vimos adotando, entendemos que os bens jurídicos protegidos são a integridade e a **disponibilidade dos sistemas informáticos**[26].

---

23 Num entendimento muito próximo do nosso, Pedro Verdelho defende que o crime de dano informático além de tutelar o património digital "pretende dar tutela à defesa da integridade e fiabilidade de dados e ao bom funcionamento dos programas informáticos" [Pedro Verdelho em anotação ao art.o 4.o da LC, in VV., Comentário das Leis Penais Extravagantes - Volume 1, p. 510 (nota 5).].

24 Em Portugal previsto no artigo 3.º da Lei do Cibercrime. Sobre o crime de falsidade informática na Lei do Cibercrime, entre outros, ler: Pedro Verdelho em anotação ao art.o 3.o, *Ibid*., p. 505-509; VENÂNCIO, Lições de Direito do Cibercrime. E da tutela penal de dados pessoais; NUNES, Os Crimes Previstos na Lei do Cibercrime, p. 67-120.

25 Previsto no artigo 5.º da Lei do Cibercrime. Sobre o crime de sabotagem informática na Lei do Cibercrime ler: Pedro Verdelho em anotação ao art. 5.o, VV., Comentário das Leis Penais Extravagantes - Volume 1, p. 512-514; VENÂNCIO, Lições de Direito do Cibercrime. E da tutela penal de dados pessoais; NUNES, Os Crimes Previstos na Lei do Cibercrime, p. 259-288.

26 Nesse sentido, Manuel David Masseno defende neste caso que o crime protege "*não só bens privados como o património e a disponibilidade do sistema informático pelo respetivo titular, mas também bens públicos como a continuidade da vida em sociedade*" [MASSENO, Manuel David; WENDT, Emerson, O RANSOMWARE

Por fim, o tipo legal de **crime de interceção ilegítima de comunicação informática**[27], contém três elementos objetivos: (1) o agente atuar não estando legitimado; e (2) fazê-lo através de meios técnicos; e (3) ter como efeito da sua atuação intercetar comunicações que se processam no interior de um sistema ou rede informáticos, a eles destinadas ou deles provenientes. Também neste tipo legal, na perspetiva subjetiva, não se exige qualquer dolo específico.

O bem jurídico protegido pelo crime de interceção ilegítima, na terminologia que vimos seguindo neste artigo, é a **confidencialidade das comunicações eletrónicas**.

Em conclusão, considerando os bens jurídicos tutelados no âmbito das medidas de cibersegurança, verificamos que: o crime de acesso ilegítimo pune a violação da confidencialidade dos sistemas informáticos; o crime de interceção ilegítima pune violação à confidencialidade das comunicações eletrónicas; o crime de dano pune as violações à integridade de dados informáticos; o crime de falsidade pune as violações à autenticidade dos dados informáticos; e o crime de sabotagem pune as violações à disponibilidade de dados e sistemas informáticos.

VI. A CRIMINALIDADE INFORMÁTICA COMO AMEAÇA À CIBERSEGURANÇA

Se os bens e valores jurídicos protegidos são os mesmos, é possível estabelecer uma correlação entre os atos que se visam prevenir na cibersegurança e os atos que se visam punir no cibercrime.

Neste particular assume especial relevância as condutas punidas pelos crimes de acesso ilegítimo a sistema informático e de interceção ilegítima de comunicações informáticas, para as quais podemos estabelecer uma correlação com as medidas de cibersegurança de identificação e prevenção e vulnerabilidades.

NA LEI : APONTAMENTOS BREVES DE DIREITO PORTUGUÊS E BRASILEIRO, Direito e TI, p. 1-13, 2017, p. 8.].

27 Previsto no artigo 7.º da Lei do Cibercrime. Sobre o crime de interceção ilegítima previsto na Lei do Cibercrime ler: Pedro Verdelho em anotação ao art.o 7.o, VV., Comentário das Leis Penais Extravagantes - Volume 1, p. 517-519; VERDELHO, Pedro, *PHISHING* e outras formas de defraudação nas redes de comunicação, Direito da Sociedade de Informação, v. VIII, p. 407 a 419, 2009; VENÂNCIO, Lições de Direito do Cibercrime. E da tutela penal de dados pessoais, p. 93-94; NUNES, Os Crimes Previstos na Lei do Cibercrime, p. 329-350.

De facto, o acesso não autorizado a sistemas informáticos ou a interceção não autorizada a comunicações eletrónicas, constituem o primeiro nível de vulnerabilidade dos sistemas a identificar e prevenir, na medida em que o acesso ao sistema (ou às comunicações eletrónicas) representa a porta de entrada para todos os demais atos maliciosos que sobre sistemas, dados ou comunicações informáticas se queiram praticar[28].

Mas, numa outra perspetiva, a forma de identificar essas vulnerabilidades de acesso a sistemas e comunicações passa, precisamente, por testar formas *incumpridoras* dos protocolos legítimos de acesso ao sistema e/ou comunicações. O que determina que os agentes externos, maliciosos ou não, que tentam aceder aos sistemas e/ou comunicações informáticas por meios tecnicamente não legítimos são, a mais das vezes, os primeiros a identificar as vulnerabilidades dos sistemas informáticos.

Existe assim uma correlação umbilical entre as práticas de cibersegurança e os atos punidos pelos crimes informáticos em sentido estrito citados. O que justifica que, embora em âmbitos distintos - na cibersegurança do lado da prevenção, no cibercrime do lado da punição - se exija uma certa harmonia entre políticas de segurança informática e políticas de combate à cibercriminalidade.

Tendo isto em vista, alguns instrumentos políticos de cibersegurança apontam linhas orientadoras também para a política legislativa em matéria de criminalidade informática, em particular, neste seu sentido estrito dos crimes que punem atos contra dados, comunicações e sistemas informáticos.

É neste contexto que no considerando (60) da Diretiva (UE) 2022/2555 do Parlamento Europeu e do Conselho, de 14 de dezembro de 2022 (conhecida como NIS2), relativa a medidas destinadas a garantir um elevado nível comum de cibersegurança na União, se pode ler:

> «*Os Estados-Membros, em cooperação com a ENISA, deverão tomar medidas para facilitar a divulgação coordenada de vulnerabilidades, definindo uma política nacional nessa matéria. No âmbito da sua política nacional, os Estados-Membros deverão procurar resolver, na medida do possível, os problemas*

28 Nesse sentido pode ler-se na jurisprudência portuguesa: «*O tipo legal de acesso ilegítimo constitui um crime de perigo abstrato e constitui uma barreira para evitar a prática e outros ilícitos mais graves*» [Excerto do sumário do Juiz António Geraldes do Acórdão do Tribunal da Relação de Guimarães de 17/11/2008, Processo n.º 233/07, in Coletâneas de Jurisprudência On-line Ref. 7922/2008.]

encontrados pelos peritos que investigam as vulnerabilidades, nomeadamente a sua potencial sujeição à responsabilidade penal, em conformidade com o direito nacional. Dado que as pessoas singulares e coletivas que investigam as vulnerabilidades podem, em alguns Estados-Membros, incorrer em responsabilidade penal e civil, ***os Estados-Membros são incentivados a adotar orientações sobre a não instauração de ações penais contra quem faz investigação no domínio da segurança da informação*** *e uma isenção de responsabilidade civil para as suas atividades.*» (negrito nosso)

Analisemos então os pontos de contacto (e fricção) entre políticas de promoção da cibersegurança e políticas de combate ao cibercrime.

VII. OS ATOS NÃO PUNÍVEIS POR INTERESSE PÚBLICO DE CIBERSEGURANÇA

Este considerando (60) da Diretiva NIS2 reflete a preocupação de que a criminalização dos acessos e interceções eletrónicas não autorizadas, ainda que sem fins maliciosos, possa constituir um entrave à partilha de informação relativa a vulnerabilidades críticas. Sugerindo, nessa medida, a despenalização destas práticas quando daí possa resultar um benefício efetivo para a cibersegurança dos sistemas eletrónicos. Querendo isto dizer, quando destes atos não resulte dano para dados, comunicações ou sistemas informáticos e se obtenha informação relevante quanto à identificação e prevenção de vulnerabilidades que, como vimos, são as duas etapas iniciais das medidas de cibersegurança.

Esta preocupação é tanto mais relevante em ordenamentos jurídicos, como o português, em que os tipos legais dos crimes de acesso ilegítimo e interceção ilegítima não preveem para a sua punibilidade um elemento subjetivo específico, como, por exemplo, a intenção de obter benefício ou causar prejuízo indevido.

No contexto português, o legislador aproveitou precisamente o processo de transposição da Diretiva (UE) 2022/2555, para alterar a Lei do Cibercrime[29] com o sentido e extensão de proceder à exclusão de punibilidade de factos suscetíveis de consubstanciar os crimes de

29 Realizado através da Lei de autorização legislativa n.º 59/2025, de 22 de outubro, e Decreto-lei autorizado n.º 125/2025, de 4 de dezembro.

acesso ilegítimo e de interceção ilegítima mediante a verificação cumulativa de um conjunto de circunstâncias que visam contribuir para garantir o elevado nível comum de cibersegurança desejado na União Europeia (artigo 5.º). Este propósito concretizou-se com o novo *artigo 8.º-A* da Lei do Cibercrime portuguesa sob a epígrafe «Atos não puníveis por interesse público de cibersegurança».

Nos números 1 a 4 do neófito artigo 8.º-A da Lei do Cibercrime expõem-se os requisitos gerais de exclusão de punibilidade das condutas integradas nos crimes de acesso ilegítimo e interceção ilegítima, com base na satisfação de um interesse público de cibersegurança. Na realidade um extenso leque de requisitos cumulativos, uns de natureza substancial e outros procedimental, que não cabe neste artigo analisar exaustivamente.

Alguns destes requisitos parecem-nos indispensáveis à satisfação do interesse público de cibersegurança, e não nos suscitam dúvidas a sua pertinência para uma norma desta natureza. É o caso de:

1) Na alínea a), um requisito subjetivo positivo, que o agente atue com a «*intenção única de identificar a existência de vulnerabilidades*» e «*contribuir para a cibersegurança*» e que não seja direta ou indiretamente responsável por tais vulnerabilidades;

2) Na alínea b), o que poderíamos designar de elemento subjetivo negativo, ou seja, a ausência por parte do agente de determinada intencionalidade maliciosa;

3) Na alínea c), que as eventuais vulnerabilidades detetadas sejam imediatamente comunicadas ao responsável pelo sistema acedido, e que, nos termos do n.º 2, sejam igualmente comunicadas ao Centro Nacional de Cibersegurança.

4) Na alínea d) que a atuação do agente seja «proporcional aos seus propósitos» - referidos na alínea a) - e «*estritamente limitada*» e «*necessárias*» para estes fins. Enumerando, exemplificativamente, três situações em que considera que a atuação não cumpre tais princípios. Para a determinação da proporcionalidade referida nesta alínea o legislador aditou ainda um conjunto de elementos a considerar no n.º 3 desta norma[30].

30 Por manifesta falta de espaço, e por não ser essencial ao raciocínio que pretendemos expor, não aprofundaremos neste texto estes requisitos.

No entanto, o legislador português optou por não incluir nesta exclusão de punibilidade os ilícitos relativos, nomeadamente, ao acesso indevido a dados pessoais. Tal facto determinou que a norma estabelecesse um complexo regime de conjugação com a tutela de dados pessoais[31] que torna a aplicabilidade prática da norma bastante diminuta, face à probabilidade do agente que promova um acesso não autorizado vir a, intencionalmente ou não, aceder a dados pessoais de terceiros.

Olhando para essa questão de forma muito resumida, começamos por salientar que o objetivo da introdução deste artigo 8.º-A da Lei do Cibercrime seria incentivar os autores de atos de acesso ou interceção ilegítimos (ou seja, não autorizados), que não tenham fins maliciosos nem causem dano aos bens protegidos, a notificar as vulnerabilidades encontradas ao seu titular do sistema e à entidade responsável pela cibersegurança. Argumenta-se que a exclusão da punibilidade destes atos, em princípio ilícitos, pode ser um fator promotor de um maior reporte de vulnerabilidades dos sistemas informáticos e uma mais célere implementação de medidas de prevenção adequadas e, com isso, contribuir para o interesse público de cibersegurança.

A criação de um regime que exclua a ilicitude ou a punibilidade de atos de intromissão não autorizada em sistemas ou comunicações informáticas quando não exista intuito malicioso, não provoque danos (aos dados, comunicações ou sistemas informáticos do titular do sistema ou de terceiros), e se limite a uma atuação proporcional e limitada aos atos necessários aos fins de deteção das vulnerabilidade existentes, é compatível (e para alguns desejável) com uma política de criação de um ecossistema digital seguro. Estes requisitos, cumulativos, encontram-se no essencial espelhados no atual artigo 8.º-A da Lei do Cibercrime.

No entanto, não admitindo o legislador português incluir nesta exclusão de punibilidade os ilícitos penais relativos à tutela de dados pessoais[32], e impondo um conjunto de deveres de comunicação a estes relativos, torna-se pouco espectável que um agente não autorizado (*hacker*) venha a voluntariamente disponibilizar a informação pertinente relativas a

31 Que pela sua complexidade e extensão também não teremos espaço de detalhar neste artigo.

32 0 que certamente suscitaria igualmente sérias questões de constitucionalidade que não analisamos nesta sede.

vulnerabilidades encontradas no sistema informático, temendo que estas incluam ou induzam a factos que possam consubstancia um ilícito contra dados pessoais[33].

O que, na prática, significará que todo o agente que cometa um acesso ilegítimo (entenda-se, não autorizado) irá sempre declarar não ter acedido a dados pessoais ou, pior, suspeitando que esse acesso a dados pessoais possa ser identificado a partir das informações comunicadas, pura e simplesmente não irá comunicar o acesso ou as vulnerabilidades encontradas. Frustrando assim o propósito da norma.

VIII. NOTAS CONCLUSIVAS

A Era da designada sociedade da informação, enquanto novo modelo de comunidade, assenta na confidencialidade, integridade, autenticidade e disponibilidades de dados, sistemas e comunicações informáticas, como elementos estruturais essenciais à preservação da sua funcionalidade enquanto comunidade social, cultural, económica e política.

As políticas de cibersegurança que preveem os procedimentos necessários a garantir estes bens e valores jurídicos, decompõe-se tradicionalmente nas medidas de: (1) identificar as vulnerabilidades dos sistemas informáticos; (2) proteger ou prevenir os riscos inerentes a estas vulnerabilidades; (3) implementação de meio de deteção de incidentes que possam explorar estas vulnerabilidades; (4) responder a estes incidentes o mais rapidamente possível; e (5) recuperar a plena funcionalidade de sistemas, dados e comunicações informáticas em caso de incidente.

Do lado da legislação de combate à cibercriminalidade, vimos que os crimes informáticos em sentido estrito harmonizados internacionalmente punem precisamente violações à confidencialidade, integridade, autenticidade e disponibilidades de dados, sistemas e comunicações informáticas.

Verificamos assim que as principais vulnerabilidades que as medidas de cibersegurança visam identificar e prevenir, nos seus dois primeiros passos, são precisamente as que permitem atos de acesso indevido ou não autorizado aos sistemas e comunicações

33 O que atinge o nível do absurdo quando ao abrigo da mesma norma estaria igualmente a confessar ter incumprido com um dos requisitos, cumulativos, para que a exclusão da punibilidade do acesso ilegítimo lhe seja reconhecida.

informáticos, ou seja, os punidos pelos crimes de acesso ilegítimo sistema informático e interceção ilegítima de comunicações eletrônicas.

Desta constatação decorre que os agentes que praticam estes atos ilícitos - ainda que sem fins maliciosos - são potencialmente detentores de informação critica sobre as vulnerabilidades dos sistemas informáticos e, por isso, relevantes para as políticas de cibersegurança.

É neste pressuposto que a introdução deste artigo 8.º-A da Lei do Cibercrime portuguesa pretende incentivar os autores de atos de acessos ilegítimos (ou seja, não autorizados) a sistemas informáticos, que não tenham fins maliciosos nem causem dano, a notificar as vulnerabilidades encontradas ao titular do sistema e à entidade responsável pela cibersegurança. Argumenta-se que a exclusão da punibilidade destes atos, em princípio ilícitos, pode ser um fator promotor de um maior reporte de vulnerabilidades essenciais às medidas de prevenção e, com isso, contribuir para o interesse público de cibersegurança.

Nesse sentido, a criação de um regime que exclua a ilicitude ou a punibilidade de atos de intromissão não autorizada em sistemas ou comunicações informáticas quando não exista intuito malicioso, não provoque danos (ao sistema, ao titular do sistema ou a terceiros), e se limite a uma atuação proporcional e limitada aos atos necessários aos fins de deteção das vulnerabilidades existentes, é compatível, e para alguns desejável, com uma política de criação de um ecossistema digital seguro.

No entanto, a solução consagrada na lei portuguesa para conjugação deste regime com as normas de tutela de dados pessoais parece-nos, no mínimo, ingénua, para não dizer, desastrosa. Acreditamos por isso que, muito provavelmente, esta incoerente conjugação com o regime de tutela de dados pessoais, irá arrastar esta causa de exclusão de punibilidade para a sua impraticabilidade.

O que justifica que se reavalie o regime proposto e se pondere alternativas mais eficazes para uma política que promova a partilha de informação relevante para as políticas de cibersegurança.

BIBLIOGRAFIA:

CNCS, Centro Nacional de Cibersegurança. **Quadro Nacional de Referência para a Cibersegurança**. [s.l.]: Gabinete Nacional de Segurança, 2019. Disponível em: <https://www.cncs.gov.pt/docs/cncs-qnrcs-2019.pdf>.

DOWNING, Douglas; COVINGTON, Michael; COVINGTON, Melody Mauldin; *et al*. **Dicionário de Termos Informáticos e da Internet**. Lisboa: Paralelo Editora, Lda, 2001.

FREITAS, Pedro. Breves nótulas sobre o crime de acesso ilegítimo previsto na Lei do Cibercrime. *In*: VV, AA (Org.). **Estudos em Comemoração dos 20 anos da Escola de Direito da Escola de Direito da Universidade do Minho**. Coimbra: Coimbra Editora, 2014, p. 565–585.

MASSENO, Manuel David; WENDT, Emerson. O RANSOMWARE NA LEI : APONTAMENTOS BREVES DE DIREITO PORTUGUÊS E BRASILEIRO. **Direito e TI**, p. 1–13, 2017.

MATOS, José António Alves de. **Dicionário de Informática e Novas Tecnologias**. LISBOA: Lidel - Edições Técnicas, L.da, 2009.

NUNES, Duarte Rodrigues. **Os Crimes Previstos na Lei do Cibercrime. E a responsabilidade penal dos entes colectivos.** 2.ª. Coimbra: GestLegal, 2024.

SIMÕES, Pedro Dias. O "Hacking" enquanto Crime de Acesso Ilegítimo. Das suas especialidades à utilização das mesmas para a fundamentação de um novo direito. **Direito da Sociedade da Informação**, v. VIII, p. 229 a 261, 2009.

VENÂNCIO, Pedro Dias. A interoperabilidade como garantia do direito fundamental à liberdade de acesso às redes informatizadas de uso público, previsto no artigo 35.º n.º 6 da C.R.P. **Anuário de Direitos Humanos**, v. 2, p. 71–88, 2019.

VENÂNCIO, Pedro Dias. **A Tutela Jurídica do Formato de Ficheiro Eletrónico**. Coimbra: Almedina, 2016.

VENÂNCIO, Pedro Dias. **Lições de Direito do Cibercrime. E da tutela penal de dados pessoais**. 2.ª. [s.l.]: Editora D'Ideias, 2025.

VENÂNCIO, Pedro Dias. **Manual de Propriedade Intelectual Digital**. 1.ª. Coimbra: Editora D'Ideias, 2023.

VERDELHO, Pedro. PHISHING e outras formas de defraudação nas redes de comunicação. **Direito da Sociedade de Informação**, v. VIII, p. 407 a 419, 2009.

VIEIRA, José Alberto. **A protecção dos programas de computador pelo Direito de Autor**. [s.l.]: Lex, 2005.

VV. **Comentário das Leis Penais Extravagantes - Volume 1**. Lisboa: Universidade Católica Portuguesa, 2010.



O NOVO ART. 8.º-A DA LEI DO CIBERCRIME: CONSIDERAÇÕES BREVES SOBRE O HACKING ÉTICO EM PORTUGAL, DESDE UMA PERSPETIVA SIMULTANEAMENTE TÉCNICA E JURÍDICA

PEDRO MARTINS PROENÇA*

&

MANUEL DAVID MASSENO †

* Docente na Unidade Politécnica Militar - Instituto Universitário Militar**.** ORCID: https://orcid.org/0009-0006-8375-7712.

† Professor Adjunto do Instituto Politécnico de Beja. ORCID: https://orcid.org/0000-0001-8861-0337.

RESUMO:

O aditamento do art. 8.º-A à Lei do Cibercrime (LC) introduziu, pela primeira vez no Ordenamento Jurídico português, um regime específico de não punibilidade aplicável a determinadas práticas de *hacking* ético, entendidas como ações destinadas a reforçar a segurança digital através da simulação autorizada de métodos utilizados por ciberatacantes. Partindo da caracterização do *hacking* ético e da evolução internacional da criminalização dos ataques contra sistemas de informação, com especial referência à Convenção de Budapeste e à recente Convenção das Nações Unidas contra o Cibercrime, é feita uma breve análise dos crimes de "acesso ilegítimo" e de "interceção ilegítima", os quais constituem o principal objeto do art. 8.º-A. Seguidamente, numa perspetiva técnico-jurídica, são analisados os pressupostos e os limites deste novo regime, bem como as principais dificuldades decorrentes da sua aplicação prática e da sua articulação com outros diplomas legais.

Palavras-chave: Acesso ilegítimo; Cibersegurança; *hacking* ético; interceção ilegítima; Lei do Cibercrime.

ABSTRACT:

The addition of Article 8-A to the Portuguese Cybercrime Law introduced, for the first time in the Portuguese legal order, a specific exemption from criminal liability applicable to certain ethical hacking practices, understood as actions aimed at strengthening digital security through the authorised simulation of methods used by cyberattackers. Starting from a characterisation of ethical hacking and the international evolution of the criminalisation of attacks against information systems, with a particular reference to the Budapest Convention and the recent United Nations Convention against Cybercrime, the paper provides a brief analysis of the offences of "illegal access" and "illegal interception", which constitute the main focus of Article 8-A. Then, from a technical and legal perspective, the requirements and limits of this new regime are subject to analysis, as well as the main difficulties arising from its practical application and its interaction with other legal instruments.

Keywords: Cybersecurity; ethical hacking; illegal access; illegal interception; Portuguese Cybercrime Law.

I. INTRODUÇÃO [1] - O *HACKING* ÉTICO

Embora as ações maliciosas remontem a já várias décadas, a pandemia COVID-19 agravou a combinação de ciberataques e cibercrimes, como acentuou logo a própria Comissão Europeia[2-3]. Neste contexto, conforme sustentam FERREIRA e GUEDES[4], os *hackers* éticos*,* ou *white hats,* surgem como especialistas na proteção de sistemas contra tais ações, redigindo programas, muitas vezes em código aberto, que ajudam na deteção e notificação de vulnerabilidades, ou seja, de "um ponto fraco, uma suscetibilidade ou uma falha de um ativo, sistema, processo ou controlo suscetível de ser explorado" (n.º 15 do art. 6.º da Diretiva NIS2[5]).

Para um melhor entendimento relativamente à caracterização do perfil profissional dos *hackers* éticos, temos que a *European Skills, Competences, Qualifications and Occupations* (ESCO)[6-7] - instrumento desenvolvido pela UE para descrever perfis profissionais e

1 O presente artigo resulta do Módulo 2 - «*Ensino de Ética em Hacking com base em Legislação Europeia, Normas Técnicas e Recomendações de Cibersegurança*», no âmbito do Projeto *Intermediate Topics in Cybersecurity* (ITCyber), cofinanciado pela Comissão Europeia (2022-1-EL01-KA220-HED-000085203), no quadro do Programa Erasmus+ Parcerias de Cooperação no Ensino Superior (KA220-HED), https://itcyber.eu/. Este Projeto integra Parceiros da Grécia, nomeadamente a *Hellenic Mediterranean University*, de Heraclião; da Polónia, com o *Professor Edward Lipinski* da *Academy of Applied Sciences* e a *Stanisław Staszic Social Initiatives Association*, ambas de Kielce; e de Portugal, com a ISU-UA - *International Sharing University* - *Universidade do Atlântico*, de Oeiras, e o *Instituto Politécnico de Beja*. Do referido Módulo, também resultou o Trabalho de Projeto de Mestrado em Engenharia de Segurança Informática do 1.º coautor, intitulado "MOOC sobre Ética em *Hacking* no âmbito do Projeto ITCyber". Em ambos, a orientação relativa à componente jurídica coube ao 2.º coautor.

2 Na sua Comunicação ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social e ao Comité das Regiões sobre a Estratégia da UE para a União da Segurança (COM(2020) 605 final), de 24 de julho de 2020, logo depois retomada na Comunicação Conjunta ao Parlamento Europeu e ao Conselho Estratégia de cibersegurança da UE para a década digital (JOIN(2020) 18 final), de 16 de dezembro de 2020, a qual acompanhou as Propostas de Regulamentos e de Diretivas apresentadas na mesma data.

3 HARJINDER SINGH LALLIE *et al.*, «Cyber security in the age of COVID-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic», *Computers & Security* 105 (2021): 102248, https://doi.org/10.1016/j.cose.2021.102248.

4 JOÃO FERREIRA e INÊS SOUSA GUEDES, *Hacking: Evolução, Perfis e Explicações Criminológicas*, *in*: GUEDES, INÊS SOUSA e GOMES, MARCUS ALAN DE MELO (Coord.), Cibercriminalidade - Novos Desafios, Ofensas e Soluções (Lisboa: PACTOR - Edições de Ciências Sociais, Forenses e da Educação, 2021), p. 189.

5 Diretiva (UE) 2022/2555 do Parlamento Europeu e do Conselho de 14 de dezembro de 2022 relativa a medidas destinadas a garantir um elevado nível comum de cibersegurança na União que altera o Regulamento (UE) n.º 910/2014 e a Diretiva (UE) 2018/1972 e revoga a Diretiva (UE) 2016/1148 (Diretiva SRI 2) (Texto relevante para efeitos do EEE).

6 Para um melhor entendimento acerca do perfil dos *hackers* éticos: ESCO, «Occupations», *Ethical Hacker*, ESCO Código 2529.4 - *ethical hacker*, https://esco.ec.europa.eu/en/classification/occupation_main.

7 Em sentido idêntico, a Agência da União Europeia para a Cibersegurança (ENISA) desenvolveu o *European Cybersecurity Skills Framework* (ECSF), um quadro europeu de perfis profissionais e competências em Cibersegurança, adotando a designação de «*penetration tester*» como título principal, em detrimento da

competências - sublinha que estes agentes realizam avaliações de vulnerabilidades e testes de intrusão (ações também designadas por *penetration testing* ou *pentesting*) de acordo com métodos reconhecidos pela indústria. Este perfil é igualmente associado a denominações com funções próximas, embora não necessariamente coincidentes, como *pentester*, *red team expert*[8], entre outras.

Além disso, os *hackers* éticos têm tido especial relevo em funções defensivas, particularmente no campo da investigação forense digital, onde utilizam ferramentas de ciberatacantes para "*recolher provas digitais valiosas, preservar a integridade das evidências e fornecer uma análise abrangente para processos legais*"[9].

Portanto, na expressão «*hacker* ético», o conceito «ético» tende a confundir-se com o cumprimento da lei. O que contrasta com a distinção geralmente aceite entre «Ética» e «lei», isto é, a Ética refere-se a princípios ou padrões para o comportamento humano baseados em valores e na definição e sistematização de conceitos como o certo, o errado e o crime[10], ao passo que a lei é imposta por instituições públicas, pelo menos em geral.

Não obstante, podem existir confrontos ético-legais no *hacking* ético, como o caso da aplicação de *patches* de segurança, que corrigem vulnerabilidades digitais suscetíveis de exploração por cibercriminosos, mas podem tornar indisponíveis os serviços durante a execução desse processo de correção[11]. Essa dualidade agrava-se em contextos onde o *patch* é executado em setores críticos como a saúde, no qual o impacto de uma indisponibilidade pode comprometer a vida humana, mas a ausência deste pode colocar em causa toda a infraestrutura informática e, consequentemente, o bem-estar ou a vida de todos os pacientes.

---

designação de «*hacker* ético». A nosso ver, esta opção terminológica é redutora, na medida em que, como constatado no ESCO, nem todas as atividades semanticamente associadas ao *hacking* ético se reconduzem à função de *penetration tester*.

8 O *red team* é um tipo de ação de Cibersegurança Ofensiva mais complexa, ampla e de duração mais prolongada comparativamente a um *pentest*. N esse sentido, ver: IVAN KOVAČEVIĆ e STJEPAN GROŠ, «*Red Teams - Pentesters, APTs, or Neither*», *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2020, pp. 1242-1249, https://doi.org/10.23919/MIPRO48935.2020.9245370.

9 D. OLADIMEJI *et al*., «*Ethical Hacking in Digital Forensics*», 2023, p. 5539, Scopus, https://doi.org/10.1109/BigData59044.2023.10386481.

10 Neste contexto específico, ESCO, «*Ethics*», https://esco.ec.europa.eu/en/classification/skills.

11 Sobre este assunto, DHIRANI *et al*. exploram a dualidade ético-legal entre o *hacking* ético e a privacidade. Ver: LUBNA LUXMI DHIRANI *et al*., «*Ethical Dilemmas and Privacy Issues in Emerging Technologies: A Review*», *Sensors (Basel, Switzerland)* 23, n.o 3 (2023): 1151, https://doi.org/10.3390/s23031151.

Então, torna-se imperativo que as equipas desenvolvam procedimentos para evitar a indisponibilidade dos seus sistemas, mesmo que seja em prol da segurança.

Contudo, apesar da distinção tradicional entre *hackers* éticos e agentes maliciosos, a linha entre a legalidade e a ilegalidade do *hacking* é ténue. O que ocorre ao ponto de um presumível *white hat* poder rapidamente transformar-se num *grey hat*, isto é, um *hacker* que opera sem intenções maliciosas (*v.g.* motivado por questões sociais), mas que, mesmo assim, viola a lei ou os princípios éticos se, designadamente, tornar uma vulnerabilidade pública, sem a explorar, por não ter obtido compensação económica pelo achado[12].

Porém, devido à proximidade material com comportamentos ilícitos de *black hats,* ou seja, de cibercriminosos, as ações de *hacking* ético têm sido controversas, gerando dúvidas a quem pretende agir em conformidade com a lei[13], que podem acentuar-se com a dispersão e fragmentação dos diplomas legais[14] e pelas sanções a que podem estar sujeitos, chegando ao ponto de muitos *hackers* evitarem reportar vulnerabilidades por receio de repercussões legais[15].

Em relação ao *hacking* ético na UE, é fundamental sublinhar que o mesmo passou a ser reconhecido como ferramenta legítima de prevenção e reforço da segurança digital,

12 DANIEL COHEN *et al.*, «*Collaboration or Separation Maximizing the Partnership between a "Gray Hat" Hacker and an Organization in a Two-Stage Cybersecurity Game*», *International Journal of Information Management Data Insights* 2, n.o 1 (2022): pp. 1-9, https://doi.org/10.1016/j.jjimei.2022.100073.

13 Neste contexto, TEMIMA HRLE *et al*. abordam as dúvidas dos utilizadores aquando da execução de ferramentas de varrimento de portas (vulgo *port scanning*), onde é possível obter os serviços utilizados por um determinado alvo e, consequentemente, verificar se existem vulnerabilidades nas versões enumeradas. Ver: TEMIMA HRLE *et al*., «*"Just a Tool, until You Stab Someone with It": Exploring Reddit Users' Questions and Advice on the Legality of Port Scans*», *Proceedings of the 2024 European Symposium on Usable Security*, ACM, 2024, pp. 322-336, https://doi.org/10.1145/3688459.3688469. Por exemplo, ainda no âmbito do *port scanning*, em 2003, na Finlândia, um adolescente foi condenado por acesso ilegítimo após utilizar uma ferramenta deste tipo (o *Nmap*), contra uma entidade bancária, com o objetivo de identificar possíveis falhas de segurança. Atualmente, a utilização destas ferramentas, por si só, não configura necessariamente o crime de acesso ilegítimo, sendo até prática comum entre profissionais de Cibersegurança no exercício das suas atividades. Contudo, quando utilizadas de forma desproporcional, podem provocar, entre outros efeitos, a degradação dos serviços e afetar a disponibilidade dos sistemas em teste. Para mais informações sobre a legalidade do varrimento de portas utilizando a ferramenta *Nmap* em https://nmap.org/book/legal-issues.html.

14 Sobre a Dispersão e Fragmentação Regulatórias no Ciberespaço, por todos, cfr: RAQUEL BRÍZIDA CASTRO, *Direito Constitucional: Ciberespaço e Tecnologia* (Coimbra: Almedina, 2023), p. 84.

15 G. LORENZINI *et al*., «It takes a pirate to know one: ethical hackers for healthcare cybersecurity», *BMC Medical Ethics* 23, n.o 1 (2022): pp. 1-8, Scopus, https://doi.org/10.1186/s12910-022-00872-y.

especialmente através da *Diretiva NIS2*, do *Regulamento DORA*[16] e de outros instrumentos atuais relacionados com a Cibersegurança da União.

Quanto a Portugal, o país tem acompanhado a evolução europeia em matéria de Cibersegurança através da adaptação da sua legislação, particularmente com a introdução do art. 8.º-A da LC, que veio legitimar determinadas práticas de *hacking* ético, embora, como analisaremos numa perspetiva técnico-legal, esse aditamento suscite dúvidas quanto ao seu alcance.

II. O *HACKING* ÉTICO NA CONVENÇÃO DE BUDAPESTE E NA CONVENÇÃO DA ONU CONTRA O CIBERCRIME

Para uma primeira aproximação a esta temática, onde o *hacking* ético e o cumprimento da lei surgem como indissociáveis, importa referir que o combate ao cibercrime pode ser entendido, num sentido lato, como uma forma de promover a Cibersegurança, na medida em que gera um efeito dissuasor relativamente a potenciais violações da segurança do Ciberespaço e dos Direitos Humanos das pessoas que nele circulam. Nesse sentido, o Conselho da Europa e a ONU têm desempenhado um papel essencial na harmonização internacional das políticas de Cibersegurança e no combate ao cibercrime, devendo também ser tidas em especial atenção as iniciativas da União Europeia, como depois veremos[17].

Em primeiro lugar, a *Convenção de Budapeste*, elaborada por iniciativa do Conselho da Europa, embora com a participação de países terceiros, como os Estados Unidos da América, o Canadá ou o Japão, é identificada como o principal instrumento jurídico internacional nesta matéria[18], influenciando diretamente a legislação europeia e nacional dos Estados-Membros, bem como de outros signatários.

16 Regulamento (UE) 2022/2554 do Parlamento Europeu e do Conselho, de 14 de dezembro de 2022, relativo à resiliência operacional digital do setor financeiro e que altera os Regulamentos (CE) n.º 1060/2009, (UE) n.º 648/2012, (UE) n.º 600/2014, (UE) n.º 909/2014 e (UE) 2016/1011 (Texto relevante para efeitos do EEE).

17 O que ocorre, sobretudo, desde a Comunicação da Comissão ao Conselho, ao Parlamento Europeu, ao Comité Económico e Social e ao Comité das Regiões - Criar uma Sociedade da Informação mais segura reforçando a segurança das infraestruturas de informação e lutando contra a cibercriminalidade (COM/2000/0890 final), de 21 de janeiro de 2001, da Decisão-Quadro 2005/222/JAI do Conselho, de 24 de fevereiro de 2005, relativa a ataques contra os sistemas de informação, e da Diretiva 2008/114/CE do Conselho, de 8 de dezembro de 2008, relativa à identificação e designação das infraestruturas críticas europeias e à avaliação da necessidade de melhorar a sua proteção (Texto relevante para efeitos do EEE).

18 CHAT LE NGUYEN e WILFRED GOLMAN, «*Diffusion of the Budapest Convention on Cybercrime and the Development of Cybercrime Legislation in Pacific Island Countries: 'Law on the Books' vs 'Law in Action'*»,

Efetivamente, do ponto de vista jurídico, o *hacking* ético tem sido desafiante pela multiplicidade de serviços de Tecnologias de Informação (TI) localizados e interconectados em diferentes ordenamentos jurídicos, mormente por força da abertura legal plasmada na transposição da *Convenção de Budapeste*. Mas também, segundo WERMUTH[19], por esta não especificar o *locus delicti*, o que faz com que a mesma ação de *hacking* possa ser legal num país e ilegal noutro. Esta omissão deliberada da Convenção torna especialmente complexa a criminalização de atividades ilícitas que pela sua natureza se dispersam por diferentes jurisdições[20].

Em segundo lugar, na mais recente *Convenção das Nações Unidas contra o Cibercrime; Reforço da Cooperação Internacional no Combate a Certos Crimes Cometidos por Meio de Sistemas de Tecnologias de Informação e Comunicação e para a Partilha de Provas em Forma Eletrónica de Crimes Graves* (doravante *Convenção de Hanói*)[21], NOVO[22] assinala a existência de riscos associados às infrações que a Convenção exige que cada Estado Parte criminalize, nomeadamente a possibilidade de investigadores de Cibersegurança que atuem de boa-fé serem alvo de processos judiciais caso acedam a sistemas sem autorização.

Computer Law & Security Review 40, n.o 105521 (2021): pp. 1-12, https://doi.org/10.1016/j.clsr.2020.105521; MANUEL DA COSTA CABRAL, *Cibersegurança e o Quadro Normativo Internacional* (Coimbra: Almedina, 2024): p. 116; assim como, entre nós, PEDRO VERDELHO «*Convenção do Cibercrime do Conselho da Europa - Comentário*», in VERDELHO, PEDRO *et al.* (Org.). Leis do Cibercrime. Vol. 1. Vila Nova de Famalicão: Centro Atlântico (2003), pp. 10-23. https://www.centroatl.pt/titulos/direito/imagens/excerto-ca-leisdocibercrime1.pdf.

19 MAIQUEL ÂNGELO WERMUTH, *Cibercriminalidade e Imigração: O Papel da Convenção de Budapeste no Enfrentamento dos Crimes de Racismo e Xenofobia*, *in*: GUEDES, INÊS SOUSA e GOMES, MARCUS ALAN DE MELO (Coord.), Cibercriminalidade - Novos Desafios, Ofensas e Soluções (Lisboa: PACTOR - Edições de Ciências Sociais, Forenses e da Educação, 2021), p. 348.

20 Exemplo disso são os ataques distribuídos de negação de serviço (*Distributed Denial-of-Service - DDoS*), nos quais um ciberatacante pode infetar sistemas geograficamente dispersos para que, em conjunto, as máquinas infetadas ataquem um determinado alvo, com o objetivo de degradar os serviços ou impedir o acesso aos mesmos.

21 ONU, «UN Cybercrime Convention - Full Text», United Nations : Office on Drugs and Crime, https://www.unodc.org/unodc/en/cybercrime/convention/text/convention-full-text.html; MANUEL DAVID MASSENO, «*Depois de "Budapeste" [?]: o processo de aprovação do Tratado e suas controvérsias*.», Aula Aberta sobre «O Novo Tratado sobre Cibercrime da ONU», aos Mestrados em Direito e Informática e em Direitos Humanos da Escola de Direito da Universidade do Minho, Braga, no dia 17 de maio de 2025, disponível aqui: https://www.academia.edu/129405540/Depois_de_Budapeste_o_processo_de_aprova%C3%A7%C3%A3o_do_Tratado_e_suas_controv%C3%A9rsias.

22 LISANDRA NOVO, «*In or Out? Managing Risks from the UN Cybercrime Convention*», *2025 17th International Conference on Cyber Conflict: The Next Step (CyCon)*, 2025, pp. 57-74, https://doi.org/10.23919/CyCon65856.2025.11103714.

Do mesmo modo, WATNEY[23] aborda as críticas dos grandes grupos tecnológicos à criminalização do acesso para efeitos de investigação, o que dificulta a distinção entre agentes maliciosos e profissionais de Cibersegurança. Complementarmente, o autor refere que a proibição prevista na interceção de transmissões de dados não públicas não permite a validação de vulnerabilidades pelos investigadores e, de forma semelhante, a manipulação de dados por parte de *pentesters* e das equipas de *red team*, no sentido de descoberta de falhas de segurança, pode ser criminalizada.

Em geral, os peritos alertam que a *Convenção de Hanói* pode ameaçar a Cibersegurança, ao permitir que os Estados obriguem profissionais de TI, bem como outros indivíduos com conhecimentos técnicos, a fornecer informações sensíveis para aceder ou apreender sistemas ou dados (n.º 4 do art. 28.º), podendo essa obrigação estender-se à divulgação forçada de vulnerabilidades, chaves de encriptação privadas ou código-fonte[24]. Ou seja, essa disposição de obrigatoriedade vai contra a atual perspetiva da UE de divulgação voluntária de vulnerabilidades (art. 29.º da Diretiva NIS2).

Assim como na *Convenção de Budapeste* (art. 6.º n.º 2 *in fine*), a referência a testes de segurança autorizados na *Convenção de Hanói* consta do n.º 2 do art. 11.º, que criminaliza o uso indevido de dispositivos. Na verdade, as semelhanças com a *Convenção de Budapeste* são substanciais, em especial no que respeita à redação dos arts. 2.º a 8.º, isto é, em relação ao Direito Penal Substantivo.

Do mesmo modo, ressaltamos que se as disposições especificamente protetoras do *hacking* ético não estavam previstas na *Convenção de Budapeste*, poderiam ter constado da *Convenção de Hanói*, o que refletiria a crescente importância que estas práticas têm vindo a desempenhar no setor. Aliás, tem sido a ONU a identificar o *pentesting* como uma importante ferramenta de gestão de risco[25]. No entanto, ainda que tais disposições não sejam consagradas no respetivo texto, é expectável que a *Convenção de Hanói* produza efeitos benéficos para

23 MURDOCH WATNEY, «*The Legal Pitfalls to Ratification of the United Nations Convention Against Cybercrime*», *Curran Associates Inc.*, The Proceedings of the 24th European Conference on Cyber Warfare and Security, ECCWS, 2025, pp. 713-719, https://doi.org/10.34190/eccws.24.1.3381.

24 NOVO «In or Out?».

25 JORGE FLORES CALLEJAS *et al.*, *Cybersecurity in the United Nations System Organizations, JIU/REP/2021/3, Report of the Joint Inspection Unit* (2021), https://www.unjiu.org/sites/www.unjiu.org/files/jiu_rep_2021_3_english.pdf.

os *hackers* éticos, na medida em que deverá contribuir para um enquadramento internacional mais homogéneo e colaborativo, reforçando a legitimidade das suas ações.

III. BREVES COMENTÁRIOS AO ACESSO ILEGÍTIMO E À INTERCEÇÃO ILEGÍTIMA EM PORTUGAL

Do levantamento das diversas definições de *hacking*, resulta que várias identificam o acesso ilegítimo como um crime-base frequentemente utilizado para enquadrar o *hacking* enquanto modelo conceptual de natureza predominantemente pejorativa, conotação que tem vindo a mudar ao longo dos anos[26].

Em relação ao acesso ilegítimo no contexto da jurisprudência portuguesa, o bem jurídico tutelado é a segurança do sistema informático, no que diz respeito ao seu controlo exclusivo[27], repercutindo-se ainda na confiança que o cidadão tem nesses sistemas[28]. Adicionalmente, de acordo com a Doutrina portuguesa dominante, esta criminalização visa a tutela do património lesado, da concorrência, da liberdade de iniciativa económica, da segurança jurídica e da privacidade[29].

Sem prejuízo da legislação anterior, no período que antecedeu a ratificação da *Convenção de Budapeste* e a transposição da Decisão Quadro n.º 2005/222/JAI, ambas pela LC[30] de 2009, Portugal exigia o dolo específico no acesso ilegítimo, por força da então vigente *Lei da*

---

26 FERREIRA e GUEDES, *Hacking: Evolução, Perfis e Explicações Criminológicas,* p. 189; bem como PEDRO SIMÕES DIAS, *O hacking enquanto crime de acesso ilegítimo. Das suas especialidades à utilização das mesmas para a fundamentação de um novo Direito*, *in* Actualidad Jurídica Uría Menéndez, 14–2006, pp. 55-72, https://www.uria.com/documentos/publicaciones/1580/documento/art04.pdf?id=2108; ou MANUEL DAVID MASSENO, «*Da criminalização do "acesso ilícito" ('Hacking') nos Ordenamentos do Brasil e de Portugal*», *in* CALHEIROS, CLARA *et al.* (Org.), 5.º Congresso de Direito na Lusofonia - Direito e novas tecnologias. Braga: Escola de Direito da Universidade do Minho / Centro de Investigação em Justiça e Governação, 2018, pp. 279-288.

27 Por exemplo, no Acórdão do Tribunal da Relação de Évora, de 9 de maio de 2023, Processo 1275/20.4JALRA.E1, relator NUNO GARCIA, o qual está disponível em https://www.dgsi.pt/jtre.nsf/134973db04f39bf2802579bf005f080b/1c5e4ff5439d5082802589b9002e2f6e.

28 PEDRO DIAS VENÂNCIO, *Lições de Direito do Cibercrime. E da tutela penal de dados pessoais*, 2.a Edição (Coimbra: Editora D'Ideias, 2025), p. 75.

29 DUARTE RODRIGUES NUNES, *Os Crimes Previstos na Lei do Cibercrime - E a Responsabilidade Penal dos Entes Coletivos*, (Coimbra: Gestlegal), 2024, p. 295.

30 Lei n.º 109/2009, de 15 de setembro, que aprova a *Lei do Cibercrime*, transpondo para a ordem jurídica interna a Decisão Quadro n.º 2005/222/JAI, do Conselho, de 24 de fevereiro, relativa a ataques contra sistemas de informação, e adapta o direito interno à Convenção sobre Cibercrime do Conselho da Europa.

Criminalidade Informática de 1991[31]. Ou seja, os agentes deviam estar conscientes de que a ação ilícita visava favorecer ou prejudicar indevidamente terceiros, para lá da intenção de obter acesso a um sistema, sendo esse um requisito para criminalizar o acesso indevido, conforme citamos abaixo.

Efetivamente, embora a *Lei da Criminalidade Informática* não contemplasse o *hacking* enquanto categoria autónoma - o qual, não raro, assenta no acesso (ou tentativa de acesso) a um sistema - a verdade é que a forma como tipificou e delimitou as condutas de acesso não autorizado podia, em termos práticos, deixar fora do âmbito do tipo certos comportamentos que, na interpretação contemporânea, tenderiam a ser reconduzidos a ações de *hacking* ético. Nessa perspetiva, NUNES[32] afirma que essas ações não eram punidas à luz do art. 7.º, porque presumir-se-ia não existir a intenção de obter um benefício ou vantagem ilegítimos.

Com a proliferação dos ciberataques, o legislador português acabou por seguir uma linha de criminalização geral, convertendo este tipo num crime de perigo, ao não exigir qualquer dano ou benefício ilegítimo para a consumação do ilícito. Nas palavras de VENÂNCIO[33], o objetivo da reformulação deste crime-meio[34] "extraordinariamente abrangente nas condutas puníveis" foi a prevenção de "crimes mais graves subsequentes ao acesso não autorizado", em oposição à legislação de alguns países onde a mera intrusão, sem perigo concreto, bem como os casos em que atos de acesso conduziram, paradoxalmente, à identificação de lacunas e fragilidades na segurança, não integraram qualquer tipo legal de crime[35].

Assim, o acesso ilegítimo passou a estar isento da intenção específica prevista na lei anterior, exigindo apenas o dolo genérico correspondente ao simples conhecimento e vontade de aceder, sem autorização, a um sistema informático - sem prejuízo dos arts. 13.º (dolo genérico e negligência) e 17.º (erro sobre a ilicitude) do *Código Penal*. Essa decisão, da qual resultou também a criminalização do *hacking* ético, refletiu-se no art. 6.º da LC, de 2009, embora tal

31 A Lei n.º 109/91, de 17 de agosto, que aprova a *Lei da Criminalidade Informática,* sobre a qual, neste particular, Dias, *O hacking enquanto crime de acesso ilegítimo*.

32 NUNES, *Os Crimes Previstos na Lei do Cibercrime - E a Responsabilidade Penal dos Entes Coletivos,* p. 292.

33 VENÂNCIO, *Lições de Direito do Cibercrime. E da tutela penal de dados pessoais*, pp. 74 e 76.

34 A título de exemplo, o acesso ilegítimo a um sistema pode servir de base para a prática de interceção ilegítima, ou seja, através de um *software* específico (*sniffers*), o atacante pode ser capaz de capturar informações de forma ilícita dentro do próprio sistema-alvo. No entanto, a interceção ilegítima também pode ser executada a partir do próprio sistema do atacante, na tentativa de capturar eventuais informações sensíveis que fossem inadvertidamente enviadas para si.

35 Cfr. Ponto 49 da Minuta do Relatório Explicativo da *Convenção de Budapeste*.

não fosse exigido pelo art. 2.º (acesso ilícito) da *Convenção de Budapeste*, em cujos termos "*qualquer uma das Partes pode exigir que ela seja cometida [...] com intenção de obter dados informáticos ou com qualquer outra intenção*"[36].

Em contrapartida, o art. 3.º (acesso ilegal a sistemas de informação) da Diretiva 2013/40/UE do Parlamento Europeu e do Conselho, de 12 de agosto de 2013, relativa a ataques contra os sistemas de informação e que substitui a Decisão-Quadro 2005/222/JAI do Conselho, não exige qualquer dolo específico, sendo uma "Diretiva de mínimos", como resulta do seu art. 1.º, em linha com o anteriormente previsto no art. 2.º (acesso ilegal aos sistemas de informação) da Decisão-Quadro[37].

Embora o artigo relativo ao acesso ilegítimo seja dotado de singularidade quanto ao tipo de crime que tipifica, determinados ilícitos, como a infeção por *ransomware*[38], podem, simultaneamente, preencher os elementos de vários tipos de crime dispostos na *Convenção*[39]. O mesmo ocorre no ordenamento jurídico português. Deste modo, estão configuradas situações de concurso real de crimes "*com outros crimes informáticos que lhe sejam subsequentes*"[40]. Por exemplo, um ciberatacante entra num sistema por meio de *ransomware* (acesso ilegítimo) e, noutro momento, interceta dados não públicos transmitidos para a vítima ou por ela (interceção ilegítima). Estas ações revelam-se igualmente como um concurso ideal, se as condutas ilícitas forem executadas num único momento[41].

---

36 A este propósito, conserva um especial interesse o texto breve de PEDRO VERDELHO, «*Anotação ao Artigo 6.o da Lei n.o 109/2009, de 15 de setembro*», *in* AAVV. (ALBUQUERQUE, PAULO PINTO & BRANCO, JOSÉ, Org.), *"Comentário das Leis Penais Extravagantes"*, Vol. 1 (Lisboa: Universidade Católica Editora, 2010): pp. 515-517.

37 Sobre esta questão, PEDRO MIGUEL FREITAS, «*Breves nótulas sobre o crime de acesso ilegítimo previsto na lei do cibercrime*», *in* AAVV. (MONTE, MÁRIO FERREIRA *et al.,* Org.), *Estudos em comemoração dos 20 anos da Escola de Direito da Universidade do Minho,* (Coimbra: Coimbra Editora, 2014): pp. 565-585.

38 O *ransomware* é um tipo de *malware* (ficheiro informático malicioso) com diferentes variantes, que normalmente permite aceder ilegitimamente a um sistema após ter sido introduzido neste (infeção), recorrendo, por vezes, a um servidor de Comando e Controlo (C&C ou C2) operado por atacantes para interagir com a máquina comprometida e, após o acesso, encripta os ficheiros ou bloqueia o sistema-alvo (destruição) para posteriormente ser exigido o pagamento de uma quantia monetária para a sua desencriptação (resgate). Para uma sua caraterização, nomeadamente, DUARTE RODRIGUES NUNES, «*O fenómeno do Ransomware e o seu enquadramento jurídico-penal*», *Cyberlaw by CIJIC*, 8 (2019), pp. 58-82 https://www.iuris.edu.pt/xms/files/Cyberlaw-by-CIJIC_8.pdf.

39 Tal como abordado pelo Comité da Convenção do Cibercrime do Conselho da Europa, sobre os *Aspetos do Ransomware Cobertos pela Convenção de Budapeste*, disponível em https://rm.coe.int/t-cy-2022-14-guidancenote-ransomware-v4adopted/1680a9355e.

40 VENÂNCIO, *Lições de Direito do Cibercrime. E da tutela penal de dados pessoais,* p. 78.

41 Designadamente, MANUEL DAVID MASSENO, «*Dos ataques de ransomware na Convenção de Budapeste sobre o Crime Cibernético, um ensaio de qualificação alternativa, desde Portugal*», *Privacy and Data*

Seguidamente, temos que o crime de interceção ilegítima, na sua forma original, encontrava-se previsto no art. 8.º da *Lei da Criminalidade Informática*. Subsequentemente, tal como sucedido no acesso ilegítimo, o crime de interceção ilegítima, agora tipificado no art. 7.º da LC, na sequência do art. 3.º da *Convenção de Budapeste*, acolheu também o art. 6.º da *Convenção*, relativo ao uso abusivo de dispositivos, o que torna a interceção um crime de perigo, "pois não requer a existência de um dano ou efetiva "interceção ilegítima" de qualquer comunicação eletrónica"[42]. Ora, para NUNES[43], tais condutas são "*passíveis de criar um perigo para a segurança das comunicações informáticas*", sem exigir a criação de um perigo efetivo, daí tratar-se de um crime de perigo abstrato.

Neste crime público, onde a tentativa é punível, o bem jurídico protegido será a inviolabilidade, a segurança, a intimidade/privacidade das comunicações[44]. Igualmente neste caso, o legislador português optou por não exigir no n.º 1 o dolo específico facultado pela *Convenção*, sem prejuízo do necessário dolo genérico[45].

Inclusive ainda mais que no acesso ilegítimo, consideramos que a interceção ilegítima também deve ser tida como um crime-meio para a consumação de outros crimes, designadamente em contextos onde é possível intercetar dados tão relevantes como as senhas de contas de utilizadores, para depois aceder, sem autorização, a essas contas.

Em relação à inclusão de "*emissões eletromagnéticas provenientes de um sistema informático que veicule esses dados*", prevista no art. 3.º da *Convenção*, o legislador português optou por não proceder (a nosso ver, corretamente) à sua especificação. Na nossa perspetiva, a interceção "*através de meios técnicos*" é suficientemente abrangente para visar não só a captura de emissões eletromagnéticas, como também a utilização de outras técnicas que não estão cingidas ao escopo eletromagnético, prevendo um cenário de constante evolução tecnológica. Tal ocorre, por exemplo, na captação maliciosa dos sons característicos produzidos pelas teclas de um computador ao serem pressionadas - chamado

---

Protection Magazine - Revista Científica na Área Jurídica, n.o 7 (2023): pp. 63-81, https://www.europeia.pt/resources/media/documents/PDPM_007.pdf.

42 VENÂNCIO, *Lições de Direito do Cibercrime. E da tutela penal de dados pessoais,* p. 94.

43 NUNES, *Os Crimes Previstos na Lei do Cibercrime - E a Responsabilidade Penal dos Entes Coletivos,* p. 333.

44 NUNES, *Os Crimes Previstos na Lei do Cibercrime - E a Responsabilidade Penal dos Entes Coletivos,* p. 331.

45 Sobre este tipo penal, são também de atender as considerações de PEDRO VERDELHO, «*Anotação ao Artigo 7.o da Lei n.o 109/2009, de 15 de setembro*», *in* AAVV. (ALBUQUERQUE, PAULO PINTO E BRANCO, JOSÉ, Org.), *"Comentário das Leis Penais Extravagantes"*, Vol. 1 (2010): pp. 518-519.

de ataque *Acoustic Side Channel* – que não constituem propriamente emissões eletromagnéticas e até podem ser mais fáceis de capturar[46]. Este tipo de ataques possibilita a dedução dos caracteres introduzidos a partir das diferenças acústicas e temporais associadas à digitação da vítima, inclusive a identificação de perfis[47].

IV. A LEGITIMAÇÃO DO *HACKING* ÉTICO EM PORTUGAL: O NOVO ART. 8.º-A DA LC

Antes de mais, cabe assinalar que, com a aprovação do Decreto-Lei n.º 125/2025[48], que transpôs a Diretiva NIS2 para o ordenamento jurídico nacional, o legislador manteve a notificação voluntária e coordenada de vulnerabilidades por qualquer pessoa singular ou coletiva, sem a imposição de obrigações adicionais (art. 45.º)[49]. Paralelamente, foram executadas na Ordem Jurídica interna as obrigações decorrentes do Regulamento Cibersegurança, bem como introduzidas alterações a diversa legislação nacional, das quais se destaca o art. 8.º-A, aditado à LC pelo art. 7.º do referido diploma.

Em termos gerais, o legislador terá procurado «retroceder» na criminalização do *hacking* ético com o intuito de identificar a existência de vulnerabilidades e em prol do «interesse público de Cibersegurança» e da segurança do Ciberespaço (al. a) do n.º 1), o que já mereceu críticas especialmente severas de NUNES[50] quanto ao respetivo alcance e à consistência técnica.

Ainda que, certamente por razões de prudência, o tenha feito depender de vários requisitos cumulativos, entre os quais: i. o agente não atuar com o propósito de obter vantagem ou promessa de vantagem económica, ficando excluída a remuneração auferida no âmbito

46 JOSHUA HARRISON *et al.*, «*A Practical Deep Learning-Based Acoustic Side Channel Attack on Keyboards*», *2023 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, 2023, pp. 270-280, https://doi.org/10.1109/EuroSPW59978.2023.00034.

47 De um modo abrangente, à análise do padrão de escrita num teclado dá-se o nome de «*Keystroke Dynamics*».

48 Decreto-Lei n.º 125/2025, de 4 de dezembro, que transpõe a Diretiva (UE) 2022/2555, relativa a medidas destinadas a garantir um elevado nível comum de cibersegurança na União.

49 Como já resultava do art. 20.º (notificação voluntária de incidentes) da Lei n.º 46/2018, de 13 de agosto, que estabelece o regime jurídico da segurança do ciberespaço, transpondo a Diretiva (UE) 2016/1148, do Parlamento Europeu e do Conselho, de 6 de julho de 2016, relativa a medidas destinadas a garantir um elevado nível comum de segurança das redes e da informação em toda a União, efetivando o previsto no também art. 20.º da Diretiva NIS1.

50 Em DUARTE RODRIGUES NUNES, «*Mit Brennender Sorge*: O Novo Artigo 8.o-A, n.o 1, da Lei do Cibercrime», *Justiça com A, 69ª Edição*, 2026, https://issuu.com/justicacoma/docs/69_edi_o_justi_a_com_a.

profissional (al. b) do n.º 1), tal como ocorre com os *hackers* éticos a trabalhar por conta de outrem; ii. o agente comunicar imediatamente as vulnerabilidades ao proprietário ou gestor do sistema, aos titulares de dados obtidos que estejam ao abrigo do *Regulamento Geral Sobre Proteção de Dados* (RGPD)[51] e das leis nacionais interconexas (al. c) do n.º 1), e ao Centro Nacional de Cibersegurança (CNCS), o qual é a autoridade nacional de Cibersegurança em Portugal (n.º 2); iii. o agente atuar de forma proporcional à finalidade de descobrir vulnerabilidades (al. d) do n.º 1), não praticando quaisquer outras ações que possam consubstanciar ilícitos penais para além do previsto nos artigos relativos ao acesso ilegítimo e à interceção ilegítima; e, ainda, iv. não violar os dados pessoais que porventura se encontrem no sistema em causa (art. 4.º, 12) do RGPD).

O último requisito elencado anteriormente configuraria a prática dos crimes de acesso indevido, de desvio de dados, de viciação ou destruição de dados, ou de inserção de dados falsos (arts. 47.º a 50.º da Lei n.º 58/2019[52]), assim como de acesso indevido aos dados, de desvio de dados, de viciação ou destruição de dados (art. 53.º, 54.º e 57.º da Lei n.º 59/2019[53]), ainda que o dever de comunicação imediata aos titulares dos dados tenha como pressuposto necessário a também exclusão da ilicitude relativamente aos crimes de acesso indevido e acesso indevido aos dados, antes referidos, além de exigir uma análise aprofundada do enquadramento da qualificação do agente enquanto «responsável pelo tratamento» (art. 4.º, 7) do RGPD) e da respetiva legitimidade no exercício de funções de interesse público (arts. 6.º, n.º 1, al. e) e 9.º, n.º 2, al. g) do RGPD, assim como o art. 3.º, n.º 1, al. j) e os arts. 5.º e 6.º da Lei n.º 59/2019), o qual ultrapassa manifestamente os objetivos deste estudo.

---

51 Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016, relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados) (Texto relevante para efeitos do EEE).

52 Lei n.º 58/2019, de 8 de agosto, que assegura a execução, na ordem jurídica nacional, do Regulamento (UE) 2016/679 do Parlamento e do Conselho, de 27 de abril de 2016, relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados. Sobre os quais ver: TIAGO GERALDO, «*Capítulo VII - Tutela administrativa e jurisdicional / Secção III - Crimes*», *in* CORDEIRO, A. BARRETO MENEZES (Coord.). *Comentário ao Regulamento Geral sobre Proteção de Dados e à Lei n.o 58/2019* (Coimbra: Almedina, 2021); VENÂNCIO, *Lições de Direito do Cibercrime. E da tutela penal de dados pessoais,* pp. 158-162.

53 Lei n.º 59/2019, de 8 de agosto, que aprova as regras relativas ao tratamento de dados pessoais para efeitos de prevenção, deteção, investigação ou repressão de infrações penais ou de execução de sanções penais, transpondo a Diretiva (UE) 2016/680 do Parlamento Europeu e do Conselho, de 27 de abril de 2016.

Ainda assim, especificamente, adiantamos que a interpretação articulada entre a al. d) e a al. e) do n.º 1, sempre do art. 8.º-A, excluirá a ilicitude sempre que a "violação de dados pessoais protegidos" quando esta resultar de uma "vulnerabilidade detetada ou da sua exploração"[54], isto é, estamos perante uma tensão normativa entre a proibição absoluta da violação de dados pessoais e a exclusão de responsabilidade em casos em que a «violação de dados pessoais», isto é, "*uma violação da segurança que provoque, de modo acidental ou ilícito, a destruição, a perda, a alteração, a divulgação ou o acesso, não autorizados, a dados pessoais transmitidos, conservados ou sujeitos a qualquer outro tipo de tratamento*" (Art. 4.º 12) do RGPD), decorra inevitavelmente da vulnerabilidade ou da exploração que foi evitada.

Em qualquer caso, o legislador pretendeu dar resposta à ausência de um enquadramento normativo específico aplicável ao *hacking* ético, circunstância que expunha estes profissionais a riscos de responsabilização, inclusivamente de natureza penal, quando procediam à exploração de vulnerabilidades, ainda que atuando com finalidades legítimas e de boa-fé[55].

Todavia, é até inexplicável que a redação do preceito não tenha circunscrito tais condutas aos profissionais de Cibersegurança, tal como preconizado pela ENISA. Com efeito, esta Agência da União Europeia recomendara aos Estados-Membros a adoção de mecanismos de qualificação e reconhecimento dos profissionais de *hacking* ético, sob pena de se acentuar a indistinção entre *white hat*s e *black hats*[56]. O que poderia ter sido alcançado através da adaptação dos requisitos técnicos do *pentester/red team* plasmados no art. 7.º do Regulamento Delegado (UE) 2025/1190[57], que especifica as características necessárias para

54 Neste particular, sublinhamos a destreza do legislador nacional ao diferenciar a «vulnerabilidade detetada» de «vulnerabilidade explorada»: existem vulnerabilidades cuja exploração depende de condições específicas do sistema-alvo (*e.g. software* que é vulnerável, mas que só é explorável numa determinada versão de um Sistema Operativo onde este está instalado).

55 GAIA FIORINELLI e MARIA VITTORIA ZUCCA, «I*s the Road to Hell Paved with Good Intentions? A Criminological and Criminal Law Analysis of Prospective Regulation for Ethical Hacking in Italy and the EU*», *ITASEC 2024: The Italian Conference on Cybersecurity*, 2024, https://ceur-ws.org/Vol-3731/paper45.pdf.

56 FIORINELLI e ZUCCA, «I*s the Road to Hell Paved with Good Intentions? A Criminological and Criminal Law Analysis of Prospective Regulation for Ethical Hacking in Italy and the EU*»; NUNES, «*Mit Brennender Sorge: O Novo Artigo 8.o-A, n.o 1, Da Lei Do Cibercrime*», pp. 51-52.

57 Regulamento Delegado (UE) 2025/1190 da Comissão, de 13 de fevereiro de 2025, que complementa o Regulamento (UE) 2022/2554 do Parlamento Europeu e do Conselho no que respeita às normas técnicas de regulamentação que especificam os critérios utilizados para identificar as entidades financeiras obrigadas a realizar testes de penetração baseados em ameaças, os requisitos e normas que regem a utilização de testadores internos, os requisitos relativos ao âmbito, metodologia e abordagem para cada uma das fases de realização dos testes, os resultados, as fases de conclusão e de correção e o tipo de cooperação em matéria de supervisão e outra cooperação relevante e necessária para a realização dos TLPT e para facilitar o reconhecimento mútuo.

a seleção de prestadores TLPT (em inglês *Threat-Led Penetration Testing*, ou Testes de Penetração Baseados em Ameaças) e os requisitos do art. 27.º do Regulamento DORA sobre o mesmo tópico, que estão alinhados com a classificação ESCO e ECSF[58].

Assim, com algum esforço interpretativo, o requisito de profissionalização dos *hackers* éticos pode ser deduzido a partir do conhecimento necessário para avaliar e garantir a proporcionalidade entre a descoberta da vulnerabilidade e o dano causado pelo acesso ilegítimo ou interceção ilegítima (n.º 3 do art. 8.º-A). Por outras palavras, para que a conduta do agente seja legitimada, este deve conseguir demonstrar que as consequências inerentes à vulnerabilidade encontrada seriam, no mínimo, tão gravosas quanto os danos causados pela sua intervenção (*e.g.* uma eventual violação da privacidade). Contudo, tal pressuposto revela-se tecnicamente inviável, uma vez que implicaria que os *hackers* éticos conseguissem antecipar a relação causa-efeito antes mesmo de conhecerem a natureza e a severidade das vulnerabilidades que venham a detetar.

Não menos preocupante é o legislador português ter somente exigido a notificação das intenções do agente após a descoberta de vulnerabilidades e não antes (al. c) do n.º 1). Destarte, o efeito inevitável será o de as equipas de Cibersegurança deixarem de ter controlo sobre ações ofensivas, legítimas e ilegítimas, que ocorrem nos seus sistemas, mesmo recorrendo a registos (*logs*) e a sistemas de monitorização para atestar a intenção dos agentes, e não só. Consequentemente, a legitimação de acessos por parte de elementos menos qualificados é suscetível de promover comportamentos negligentes, pondo em causa a proteção dos bens jurídicos tutelados.

Posto isto, em matéria de proporcionalidade, é recomendável que a atuação se circunscreva ao estritamente necessário para evidenciar a vulnerabilidade em causa, devendo ser evitada a prossecução de ações de *hacking* de forma consecutiva. Dito de outro modo, uma vez obtido o acesso a um sistema, o *hacker* ético deverá cessar de imediato a sua intervenção, abstendo-se de explorar, pesquisar ou testar outras vulnerabilidades no seu interior, ainda que com a

58 Deixamos a nota de que a Lei n.º 73/2025, de 23 de dezembro, que assegura a implementação de atos jurídicos europeus no ordenamento jurídico nacional relativos à resiliência operacional digital do setor financeiro, impõe a aplicação de uma coima às entidades que não cumprirem os deveres relativos à "*recolha de informações e à realização de avaliações sobre vulnerabilidades e incidentes relacionados com as TIC*" (al. xxix), da al. c), do n.º 1 do art. 10.º).

finalidade de as identificar, na medida em que tais vulnerabilidades apenas se tornam detetáveis em virtude do acesso inicial (exceto se estiver autorizado).

Acresce que a avaliação da proporcionalidade deve igualmente considerar o efeito cumulativo de atuações de outros agentes, quando tal seja previsível, porque a soma de múltiplas explorações, ainda que individualmente «moderadas», pode, no conjunto, degradar o serviço ou aumentar significativamente o risco para dados e sistemas.

No contexto das práticas de *hacking* proibidas descritas no n.º 3, importa salientar a menção à utilização de engenharia social, de *phishing* e das suas variantes, que se enquadrará na previsão típica do crime de falsidade informática da LC (art. 3.º). Aliás, estas técnicas surgem de forma redundante nas alíneas b) e c), uma vez que o *phishing* constitui, por definição, um tipo específico de ataque de engenharia social, pelo que a sua separação normativa evidencia, a nosso ver, mais uma ênfase do que uma distinção conceptual.

Quanto à "alteração dolosa de dados informáticos", prevista na al. e) do n.º 3, o legislador limitou a tipificação às ações que alterem dados sem autorização, bastando para a configuração do crime que o agente atue com dolo genérico, ou seja, com a intenção de modificar os dados sem consentimento do legítimo proprietário, o que corresponde aos crimes de dano relativo a programas ou outros dados informáticos e de sabotagem informática, previstos e punidos pelos arts. 4.º e 5.º da LC, respetivamente. Essa condicionante aplica-se no contexto de testes destinados à mitigação de ataques do tipo *Cross-Site Scripting* (XSS) – uma técnica maliciosa que possibilita a inserção ou alteração não autorizada de conteúdos em aplicações *web*, mediante a injeção de código.

Na prática, o código injetado pelo XSS pode comprometer os dados armazenados ou processados pelos clientes e servidores, podendo, até, culminar com o acesso não autorizado às máquinas dos próprios clientes após o acesso à aplicação *web* infetada. Em relação ao *defacement* – ciberataque realizado com o "*objetivo de modificar ou danificar o conteúdo ou a estética da página de um site da Internet, com o intuito de degradar ou desmoralizar as informações transmitidas*"[59] - o mesmo pode ser também executado recorrendo aos fundamentos técnicos do XSS.

59 Acórdão do Supremo Tribunal de Justiça, de 23 de junho de 2022, Processo n.º 225/20.2TELSB.S1, relatora HELENA MONIZ, disponível em https://diariodarepublica.pt/dr/detalhe/acordao/225-2022-188266275.

Todavia, testar vulnerabilidades como o XSS, ainda que em contexto de *hacking* ético e para efeitos de prova de conceito, pode acarretar consequências jurídicas. Desde logo, a alteração do aspeto de uma aplicação *web* pode interferir com a sua normal capacidade de uso, sobretudo se a estética constituir um elemento essencial à experiência do utilizador. Caso tais alterações sejam visíveis a terceiros, poderá ficar comprometida a segurança e o prestígio da página, na medida em que a vulnerabilidade é exposta publicamente. Desta forma, é desaconselhável todo e qualquer teste de segurança que implique necessariamente a introdução não autorizada de dados no sistema, mesmo quando para efeitos de prova de conceito.

Outro dos pontos de destaque no n.º 3 é a ambiguidade da proibição de "instalação e distribuição de *software* malicioso" (al. g) do n.º 3). Não tendo o legislador definido o conceito de «*software* malicioso», quer na *Diretiva NIS2*, quer no diploma de transposição, assumimos que está interdita a utilização dos já tradicionais tipos de *malware* (*trojans*, vírus informáticos, *worms, spyware, webshells*). Similarmente à nossa visão sobre a al. e), consideramos que as técnicas que insiram ficheiros maliciosos no sistema-alvo para permitir o acesso remoto - como as usadas nos designados ataques de *Local File Inclusion*[60] - devem ser evitadas por dois motivos: o primeiro, pela desproporção da prossecução de ações de *hacking* de forma consecutiva, que já referimos anteriormente, ou seja, a descoberta da vulnerabilidade e consequente acesso ao sistema para descoberta de novas vulnerabilidades, havendo o risco de colocar em causa o bem tutelado; o segundo, pelo facto de alguns ficheiros poderem ser encarados pelos sistemas de defesa como *malware,* podendo ainda ser explorados por outros agentes, o que vai contra as práticas autorizadas.

Por vezes, a exploração controlada de vulnerabilidades constitui a única via disponível para confirmar a existência da vulnerabilidade em causa. Nesses casos, compete ao *hacker* ético avaliar se a exploração se revela necessária e proporcional aos fins legítimos que se pretendem atingir ou se, pelo contrário, o simples reporte fundamentado de uma suspeita de vulnerabilidade se mostra suficiente (aconselhamos esta via, se existirem fortes indícios), evitando-se, assim, uma ingerência adicional nos sistemas aparentemente vulneráveis.

60 É uma vulnerabilidade em aplicações *web* que permite a um atacante induzir o servidor a ler ou executar arquivos locais. Determinados ficheiros maliciosos também podem ser enviados para o alvo para serem executados a partir desta técnica.

Em relação ao reporte de vulnerabilidades por parte de pessoas singulares e coletivas, o Direito da UE determina que os Estados-Membros assegurem a possibilidade de comunicação anónima às CSIRT, se tal for solicitado (n.º 1 do art. 12.º da *Diretiva NIS2*). Esta opção visa promover a divulgação responsável de vulnerabilidades, permitindo o reporte sem recolha de elementos identificativos desde a origem.

Em contraste, o regime português estabelece que o CERT.PT - *Computer Emergency Response Team*, que pertence ao CNCS e que integra a rede CSIRT europeia – preserva o anonimato do comunicante, caso este o solicite. Ora, esta formulação pressupõe a eventual identificação prévia da pessoa singular ou coletiva e que se traduz num dever de confidencialidade posterior, e não necessariamente num mecanismo de reporte anónimo, conforme o n.º 4 do art. 38.º do Anexo a que se refere o art. 2.º do RJC. A proteção do anonimato surge ainda condicionada pelo disposto na LC, uma vez que a ocultação da identidade não pode funcionar como um mecanismo de salvaguarda para a prática de crimes informáticos, nem comprometer o princípio da responsabilidade dos *hackers* éticos que o novo diploma de transposição pretende alcançar.

Antes de terminar este breve excurso, cumpre ainda evidenciar o âmbito muito alargado do novo art. 8.º-A da LC. Concretamente, embora a *Diretiva NIS2* se aplique exclusivamente a entidades essenciais ou importantes, o regime de não punibilidade consagrado no art. 8.º-A da LC assume um alcance geral, abrangendo a descoberta de vulnerabilidades, independentemente de o respetivo titular se encontrar ou não sujeito a esta Diretiva. A decisão do legislador português sobre essa matéria poderá ser justificada pelo facto de qualquer serviço colocado em acesso público, mesmo que seja para fins pessoais (*e.g.* um *website* publicado na Internet), poder ser alvo de ciberataques que vão utilizar eventuais vulnerabilidades encontradas para realizar novos ataques a terceiros (*e.g.* a vulnerabilidade XSS num *website*, gerido por uma pessoa singular, pode infetar os utilizadores que visitem essa página)[61].

61 Em sentido análogo, por exemplo, o Tribunal de Justiça concluiu que a videovigilância que abrange o espaço público, realizada por um particular, não constitui uma atividade exclusivamente pessoal ou doméstica, ficando sujeita ao RGPD. Ver Acórdão do Tribunal de Justiça (Quarta Secção), de 14 de dezembro de 2014, Processo C-212/13, FRANTIŠEK RYNEŠ, ver https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=celex%3A62013CJ0212.

Simultaneamente, uma tal decisão levanta preocupações relativamente à proteção dos direitos fundamentais, pois o "interesse público de Cibersegurança" é ambíguo e, assim, qualquer privado, pessoa singular ou coletiva, pode ficar sob o escrutínio legitimado dos *hackers* éticos e não éticos.

O que ainda é mais preocupante no que se refere à exclusão da ilicitude no referente ao crime de interceção ilegítima. De facto, como bem acentua NUNES, o preceito em análise permitiria "*a devassa do sigilo das comunicações privadas, sendo que, nos termos do art. 34.º n.ºs 1 e 4, da CRP, o sigilo da correspondência e dos outros meios de comunicação privada é inviolável e é proibida toda ingerência das autoridades públicas na correspondência, nas telecomunicações e nos demais meios de comunicação, salvo os casos previstos na Lei em matéria de processo criminal*". Sendo impensável, atendendo às garantias consagradas na Lei Fundamental que essa devassa fosse admitida por um particular e sem autorização judicial prévia, por muito éticas que fossem as respetivas intenções[62].

Por outro lado, as referências expressas às Leis n.º 58/2019 e 59/2019, pelo art. 8.º-A (al. e) do n.º 1), alargam o alcance da exclusão da ilicitude das ações de *hacking* ético aos Órgãos de Polícia Criminal, ao Ministério Público e aos Tribunais Criminais, mesmo estando estas entidades excluídas do âmbito de aplicação subjetivo do regime jurídico da Cibersegurança (art. 3.º n.º 7 b) e e) do RJC).

Em suma, este regime de exclusão da ilicitude apenas não alcançará os sistemas informáticos do "*Estado-Maior-General das Forças Armadas e dos ramos das Forças Armadas, no que respeita às redes e sistemas de informação diretamente relacionados com o seu comando e controlo*", "as entidades públicas com responsabilidades exclusivas em matéria de produção de informações, nomeadamente ao Sistema de Informações da República Portuguesa, ao Serviço de Informações Estratégicas de Defesa e ao Serviço de Informações de Segurança, no que respeita às redes e sistemas de informação diretamente relacionados com o seu comando e controlo" e as "entidades públicas cuja atividade incida sobre redes e sistemas de informação diretamente relacionados com a produção e difusão de informação classificada, nomeadamente com as marcas nacionais, da Organização do Tratado do Atlântico Norte (OTAN), e da União Europeia, ou catalogada como segredo de Estado, no que respeita a essas redes e sistemas de informação." (art. 3.º n.º 7 als. a), c) e d) do RJC), em consonância

62 Assim, NUNES, «*Mit Brennender Sorge: O Novo Artigo 8.o-A, n.o 1, da Lei do Cibercrime*», pp. 53-54.

com as disciplinas relativas à proteção de dados pessoais (art. 2.º, n.º 2, al. b) do RGPD e art. 2.º, n.º 3, da Lei n.º 59/2019).

V. NOTAS CONCLUSIVAS

Ainda que subsistam dúvidas quanto à substância do art. 8.º-A e algumas reservas quanto ao alcance efetivo de um diploma que procura salvaguardar o denominado "interesse público de Cibersegurança", não deixa de ser um marco relevante na legitimação do *hacking* ético em Portugal. Nessa medida, não se pode prescindir nem do *hacking* ético enquanto prática essencial à identificação e mitigação de vulnerabilidades, nem desta abertura legal que, embora limitada, contribui para a integração do *hacking* num quadro jurídico compatível com as exigências atuais da Cibersegurança e da UE.

Em jeito de conclusão, espera-se que a próxima *Estratégia Nacional de Segurança do Ciberespaço* em Portugal, sem atualização desde 2023, venha reforçar não apenas as preocupações legais, mas também as questões técnicas associadas às práticas de *hacking* ético. Até lá, uma vez que diplomas legais como o RGPD apenas enunciam padrões genéricos e não vinculativos de normas de segurança para assegurar a proteção e a resiliência dos sistemas e dos serviços de tratamento, sugerimos que os *hackers* éticos que atuam no Ordenamento jurídico português e da União se guiem pelas normas técnicas mais relevantes na área (*e.g. Open Worldwide Application Security Project*, *The Penetration Testing Execution Standard*, etc.), bem como por um código de conduta, conforme o esquema autorregulatório previsto nos arts. 40.º e 41.º do RGPD.

BIBLIOGRAFIA

CABRAL, MANUEL DA COSTA. Cibersegurança e o Quadro Normativo Internacional. Coimbra: Almedina, 2024.

CASTRO, RAQUEL BRÍZIDA. Direito Constitucional: Ciberespaço e Tecnologia. Coimbra: Almedina, 2023.

COHEN, DANIEL, AMIR ELALOUF, e RAZ ZEEV. «*Collaboration or Separation Maximizing the Partnership between a "Gray Hat" Hacker and an Organization in a Two-Stage Cybersecurity Game*». International Journal of Information Management Data Insights 2, n.º 1 (2022): pp. 1-9. https://doi.org/10.1016/j.jjimei.2022.100073.

DHIRANI, LUBNA LUXMI, NOORAIN MUKHTIAR, BHAWANI SHANKAR CHOWDHRY, e THOMAS NEWE. «*Ethical Dilemmas and Privacy Issues in Emerging Technologies: A Review*». Sensors, Vol. 23, n.º 3 (2023): 1151. https://doi.org/10.3390/s23031151.

DIAS, PEDRO SIMÕES. «*O hacking enquanto crime de acesso ilegítimo. Das suas especialidades à utilização das mesmas para a fundamentação de um novo Direito*». Actualidad Jurídica Uría Menéndez, n.º 14 (2006), pp. 55-72.

ESCO. «Ethics». https://esco.ec.europa.eu/en/classification/skills.

ESCO. «Occupations». Ethical Hacker. https://esco.ec.europa.eu/en/classification/occupation_main.

FERREIRA, JOÃO e INÊS SOUSA GUEDES. Hacking: Evolução, Perfis e Explicações Criminológicas, *in*: GUEDES, INÊS SOUSA e GOMES, MARCUS ALAN DE MELO (Coord.), Cibercriminalidade - Novos Desafios, Ofensas e Soluções. Lisboa: PACTOR - Edições de Ciências Sociais, Forenses e da Educação, 2021, pp. 181-201.

FIORINELLI, GAIA, e MARIA VITTORIA ZUCCA. «*Is the Road to Hell Paved with Good Intentions? A Criminological and Criminal Law Analysis of Prospective Regulation for Ethical Hacking in Italy and the EU*». ITASEC 2024: The Italian Conference on Cybersecurity, 2024. https://ceur-ws.org/Vol-3731/paper45.pdf.

FREITAS, PEDRO MIGUEL. «*Breves nótulas sobre o crime de acesso ilegítimo previsto na lei do cibercrime*», *in* MONTE, MÁRIO FERREIRA *et al.* (Org.). Estudos em comemoração dos 20 anos da Escola de Direito da Universidade do Minho. Coimbra: Coimbra Editora, 2014, pp. 565-585.

GERALDO, TIAGO. «*Capítulo VII - Tutela administrativa e jurisdicional / Secção III - Crimes*». *In* CORDEIRO, A. BARRETO MENEZES (Coord.). Comentário ao Regulamento Geral sobre Proteção de Dados e à Lei n.º 58/2019. Coimbra: Almedina, 2021.

HARRISON, JOSHUA, EHSAN TOREINI, e MARYAM MEHRNEZHAD. «*A Practical Deep Learning-Based Acoustic Side Channel Attack on Keyboards*». 2023 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), 2023, pp. 270-280. https://doi.org/10.1109/EuroSPW59978.2023.00034.

HRLE, TEMIMA, MARY MILAD, JINGJIE LI, e DANIEL WOODS. *«"Just a Tool, until You Stab Someone with It": Exploring Reddit Users' Questions and Advice on the Legality of Port Scans*». Proceedings of the 2024 European Symposium on Usable Security, 2024, pp. 322–336. https://doi.org/10.1145/3688459.3688469.

KOVAČEVIĆ, IVAN, e STJEPAN GROŠ. «Red Teams - Pentesters, APTs, or Neither». 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), 2020, pp. 1242-1249. https://doi.org/10.23919/MIPRO48935.2020.9245370.

LALLIE, HARJINDER SINGH, LYNSAY A. SHEPHERD, JASON R. C. NURSE, *et al*. «*Cyber security in the age of COVID-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic*». Computers & Security 105 (2021): 102248. https://doi.org/10.1016/j.cose.2021.102248.

LORENZINI, G., D. M. SHAW, e B. S. ELGER. «It takes a pirate to know one: ethical hackers for healthcare cybersecurity». BMC Medical Ethics 23, n.º 1 (2022): pp. 1-8. Scopus. https://doi.org/10.1186/s12910-022-00872-y.

MASSENO, MANUEL DAVID. «*Da criminalização do "acesso ilícito" ('Hacking') nos Ordenamentos do Brasil e de Portugal*», *in* CALHEIROS, CLARA *et al.* (Org.), 5.º Congresso de Direito na Lusofonia - Direito e novas tecnologias. Braga: Escola de Direito da Universidade do Minho / Centro de Investigação em Justiça e Governação (2018), pp. 279-288.

MASSENO, MANUEL DAVID. «*Depois de "Budapeste" [?]: o processo de aprovação do Tratado e suas controvérsias.» Aula Aberta sobre «O Novo Tratado sobre Cibercrime da ONU*», aos Mestrados em Direito e Informática e em Direitos Humanos da Universidade do Minho, Braga, 17 de maio, 2025.

https://www.academia.edu/129405540/Depois_de_Budapeste_o_processo_de_aprovação_do_Tratado_e_suas_controvérsias.

MASSENO, MANUEL DAVID. «*Dos ataques de ransomware na Convenção de Budapeste sobre o Crime Cibernético, um ensaio de qualificação alternativa, desde Portugal*». Privacy and Data Protection Magazine - Revista Científica na Área Jurídica, n.º 7 (2023): pp. 63-81. https://www.europeia.pt/resources/media/documents/PDPM_007.pdf.

NGUYEN, CHAT LE, e WILFRED GOLMAN. «*Diffusion of the Budapest Convention on Cybercrime and the Development of Cybercrime Legislation in Pacific Island Countries: 'Law on the Books' vs 'Law in Action'*». Computer Law & Security Review 40, n.º 105521 (2021): pp. 1-12. https://doi.org/10.1016/j.clsr.2020.105521.

NOVO, LISANDRA. «*In or Out? Managing Risks from the UN Cybercrime Convention*». 2025 17th International Conference on Cyber Conflict: The Next Step (CyCon), 2025, pp. 57-74. https://doi.org/10.23919/CyCon65856.2025.11103714.

NUNES, DUARTE RODRIGUES. «*O fenómeno do Ransomware e o seu enquadramento jurídico-penal*», *Cyberlaw by CIJIC*, n.º 8 (2019), pp. 58-82 https://www.iuris.edu.pt/xms/files/Cyberlaw-by-CIJIC_8.pdf.

NUNES, DUARTE RODRIGUES. Os Crimes Previstos na Lei do Cibercrime - E a Responsabilidade Penal dos Entes Coletivos. Coimbra: Gestlegal, 2024.

NUNES, DUARTE RODRIGUES. «*Mit Brennender Sorge: O Novo Artigo 8.º-A, n.º 1, da Lei do Cibercrime*». Justiça com A, 69ª Edição, 2026. https://issuu.com/justicacoma/docs/69_edi_o_justi_a_com_a.

OLADIMEJI, D., L. GARLAND, J. RAMAMOORTHY, e B. ZHOU. «*Ethical Hacking in Digital Forensics*». 2023, pp. 5539-5548. Scopus. https://doi.org/10.1109/BigData59044.2023.10386481.

ONU. «*UN Cybercrime Convention - Full Text*». United Nations : Office on Drugs and Crime. https://www.unodc.org/unodc/en/cybercrime/convention/text/convention-full-text.html.

VENÂNCIO, PEDRO DIAS. Lições de Direito do Cibercrime. E da tutela penal de dados pessoais. 2.ª Edição. Coimbra: Editora D'Ideias, 2025.

VERDELHO, PEDRO. «*Convenção do Cibercrime do Conselho da Europa - Comentário*», *in* VERDELHO, PEDRO *et al.* (Org.). Leis do Cibercrime. Vol. 1. Vila Nova de Famalicão:

Centro Atlântico (2003), pp. 10-23. https://www.centroatl.pt/titulos/direito/imagens/excerto-ca-leisdocibercrime1.pdf.

VERDELHO, PEDRO. «*Anotação ao Artigo 6.º da Lei n.º 109/2009, de 15 de setembro*», *in* ALBUQUERQUE, PAULO PINTO e BRANCO, JOSÉ (Org.), Comentário das Leis Penais Extravagantes, Vol. 1. Lisboa: Universidade Católica Editora (2010), pp. 515-517.

VERDELHO, PEDRO. «*A Convenção do Cibercrime do Conselho da Europa - Comentário Anotação ao Artigo 7.º da Lei n.º 109/2009, de 15 de setembro*», *in* ALBUQUERQUE, PAULO PINTO e BRANCO, JOSÉ (Org.), Comentário das Leis Penais Extravagantes, Vol. 1. Lisboa: Universidade Católica Editora (2010), pp. 518-519.

WATNEY, MURDOCH. «*The Legal Pitfalls to Ratification of the United Nations Convention Against Cybercrime*». Curran Associates Inc., The Proceedings of the 24th European Conference on Cyber Warfare and Security, ECCWS, 2025, pp. 713-719. https://doi.org/10.34190/eccws.24.1.3381.

WERMUTH, MAIQUEL ÂNGELO. Cibercriminalidade e Imigração: O Papel da Convenção de Budapeste no Enfrentamento dos Crimes de Racismo e Xenofobia, *in:* GUEDES, INÊS SOUSA e GOMES, MARCUS ALAN DE MELO (Coord.), Cibercriminalidade - Novos Desafios, Ofensas e Soluções. Lisboa: PACTOR - Edições de Ciências Sociais, Forenses e da Educação, 2021, pp. 343-358.



O *CYBERCRIME-AS-A-SERVICE* (*CAAS*)

DUARTE RODRIGUES NUNES *

*Professor da Faculdade de Direito de Lisboa e da Universidade Europeia. Doutor em Direito. Jurisconsulto. Investigador do CIDPCC, CIJIC e *Privacy and Data Protection Centre.* Juiz de Direito em situação de licença sem vencimento. Advogado com inscrição suspensa..

RESUMO:

O *Cybercrime-as-a-Service* (*CaaS*) representa uma nova etapa na evolução do cibercrime, onde ferramentas e serviços maliciosos são disponibilizados a uma ampla gama de utilizadores em mercados e fóruns da *Dark Web* e em serviços de comunicações encriptadas. O *CaaS* democratizou o cibercrime, permitindo que criminosos com poucas aptidões na área da informática pratiquem fraudes online e executem ciberataques, como ataques de *DDoS*, campanhas de *Phishing* ou de *Cryptojacking* e ataques de *Ransomware* com facilidade. Este artigo explora o fenómeno do *CaaS*, incluindo o uso de criptomoedas para transações e o alcance global dessas plataformas ilícitas, as características dos vários intervenientes no *CaaS* e as ligações entre o *CaaS* e outras formas de criminalidade, como o terrorismo, o crime organizado, o crime económico, a ciberguerra ilícita e a ciberespionagem. Pelas suas características, o *CaaS*, além de causar enormes prejuízos financeiros, constitui uma ameaça para a segurança nacional, a segurança dos cidadãos, a estabilidade do sistema económico e financeiro e a privacidade das pessoas. Finalmente, analisa-se a resposta ao *CaaS* na ordem jurídica portuguesa e identificam-se aspetos que carecem de melhoria.

Palavras-chave: Cibercrime; *Cybercrime-as-a-Service*; Criptomoedas; *Malware*; Cibercrime organizado.

ABSTRACT:

Cybercrime-as-a-Service (*CaaS*) represents a new stage in the evolution of cybercrime, where malicious tools and services are made available to a wide range of users on Dark Web marketplaces and forums, and on encrypted communication services. *CaaS* has democratized cybercrime, allowing criminals with little computer skills to easily execute online fraud and cyberattacks, such as DDoS attacks, phishing or cryptojacking campaigns, and ransomware attacks. This article explores the phenomenon of *CaaS*, including the use of cryptocurrencies for transactions and the global reach of those illicit platforms, the characteristics of the various actors in *CaaS*, and the links between *CaaS* and other forms of crime, such as terrorism, organised crime, economic crime, illicit cyberwarfare, and cyber espionage. Due to its characteristics, *CaaS*, in addition to causing enormous financial losses, constitutes a threat to national security, the safety of citizens, the stability of the economic and financial system, and the privacy of citizens. Finally, the response to *CaaS* in the Portuguese legal system is analysed, and aspects that require improvement are identified.

Keywords: Cybercrime; Cybercrime-as-a-Service; Cryptocurrency; Malware; Organized Cybercrime.

I. INTRODUÇÃO

Nos últimos anos, assistimos ao surgimento de uma nova realidade criminosa, que, como veremos, constitui uma nova etapa evolução do cibercrime: o *Cybercrime-as-a-Service* (*CaaS*).

O cibercrime consiste na prática de factos tipificados na Lei como crime em que o sistema informático ou os dados informáticos são objeto ou instrumentos do crime ou cujo cometimento está significativamente ligado à utilização de um sistema informático e/ou de dados informáticos (Duarte Rodrigues Nunes, 2024, 66).

Por sua banda, o *CaaS* consiste na disponibilização de ferramentas e/ou serviços destinados à prática de crimes informáticos a qualquer pessoa disposta a pagar por eles. Como facilmente se percebe, o *CaaS* revolucionou o cibercrime ao proporcionar a indivíduos com menores aptidões em matéria de informática meios que lhes permitem cometer crimes informáticos (incluindo ataques informáticos complexos e altamente danosos) que, de outro modo, não conseguiriam cometer, aumentando assim, de forma drástica, o número de cibercriminosos e de potenciais cibercriminosos (Ganguli, 2024, pp. 1-2; Europol, 2023a, p. 20).

Como resultado da disponibilização destas ferramentas e serviços a qualquer pessoa disposta a pagar por eles, o cibercrime deixou de ser um exclusivo de *hackers*[1] altamente capacitados em termos de informática (pois a criação das ferramentas maliciosas e a sua utilização requeria a posse de capacidades técnicas que a generalidade dos utilizadores da informática não possui). Por isso, o cibercrime passou a ser praticado por uma miríade de outros criminosos que, de outro modo, não teriam capacidade para praticar crimes informáticos ou,

1 O *Hacking* consiste em atividades ilícitas de acesso e penetração em sistemas informáticos alheios. No âmbito do *Hacking*, costuma distinguir-se entre *Hacking* não ético (aquele que visa provocar prejuízos à vítima ou obter benefícios para o agente ou para terceiro) e *Hacking* ético (aquele em que não são provocados quaisquer prejuízos, podendo existir, inclusivamente, uma cooperação com o detentor do sistema visado, avisando-o de eventuais falhas de segurança). O agente do *Hacking* ético é o *hacker*, ao passo que o agente do *Hacking* não ético pode ser um *cracker* (aquele que pratica o *Hacking* com a finalidade de aceder ao sistema informático alheio para aí causar prejuízos, como, por exemplo, roubo de informações, obtenção de *passwords* e outros códigos de acesso, espionagem, entravamento do sistema, destruição de dados), um *phreaker* (aquele que pratica o *Hacking* acedendo e manipulando redes de telecomunicações ligadas à rede pública, com o objetivo de copiar cartões, realizar interceções de comunicações, realizar chamadas telefónicas sem pagar, usufruir de serviços de televisão sem pagar, etc.) ou um *carder* (aquele que pratica o *Hacking* com a finalidade de aceder ao sistema informático alheio para roubar códigos de acesso e códigos de cartões bancários para realizar compras *online,* levantamentos de dinheiro em caixas automáticas, transferências de dinheiro, etc.) [DUARTE RODRIGUES NUNES, 2024, p. 292 (nota 435)].

pelo menos, para praticar crimes informáticos com um tal grau de danosidade/perigosidade e complexidade.

Além disso, o *CaaS* tornou-se num modelo de negócio similar ao negócio legítimo do *Software-as-a-Service*[2], dele se distinguido pelo facto de as ferramentas e/ou os serviços disponibilizados no âmbito do *CaaS* serem destinados a permitir/facilitar a ulterior prática de crimes (Ganguli, 2024, p. 2; Europol, 2024, p. 8; Tatiana Tropina, 2013, pp. 50-51, referindo-se ao cibercrime em geral).

À semelhança do que sucede com outros mercados ilícitos *online* (*v.g.*, drogas, conteúdos de pornografia infantil, armas, bens lícitos contrabandeados, exploração de prostituição, etc.), o *CaaS* tira partido do uso de meios informáticos para dificultar a deteção dos crimes e a investigação criminal e facilitar a preparação e a execução dos crimes e o apagamento dos seus vestígios, constituindo uma atividade altamente lucrativa e que acarreta riscos penais[3] muito baixos para os criminosos (Ganguli, 2024, p. 2), sendo que os cibercriminosos, enquanto criminosos altamente racionais, tendem a ponderar os riscos penais em que incorrem face aos benefícios que poderão obter ao cometer crimes informáticos (Europol, 2021, p. 20) . No fundo, o *CaaS*, enquanto modalidade de cibercrime, proporciona aos respetivos agentes as vantagens que o uso de meios informáticos oferece a quem os utiliza na prossecução da sua atividade criminosa[4].

Além de acarretar riscos penais muito pouco substanciais, o *CaaS* é uma atividade criminosa extremamente lucrativa, sendo os lucros obtidos usufruídos pelo agente do crime (que pode ser uma ou mais pessoas físicas agindo de uma forma "não organizada", um ente coletivo, uma organização criminosa ou terrorista ou um Estado), reinvestidos em novas atividades criminosas e/ou utilizados para corromper agentes públicos ou para pagar a colaboradores externos de organizações criminosas; no caso de Estados (*v.g.*, a Coreia do Norte), o lucro é utilizado para financiar programas de armamento ou operações ilícitas e, no caso de organizações terroristas e de outros grupos subversivos, o lucro será utilizado para financiar a atividade terrorista ou subversiva.

---

2 O *Software-as-a-Service* (SaaS) consiste na disponibilização de *software* armazenado na nuvem, em que os utilizadores, em vez de comprarem e instalarem o *software* nos seus sistemas informáticos, subscrevem esse serviço, pagando um determinado preço, e acedem e utilizam esse *software* através da Internet.

3 Os "riscos penais" englobam (1) as probabilidades de punição e (2) o grau de severidade das consequências jurídicas em que o agente poderá incorrer (sanções penais ou de outra natureza e confisco de bens).

4 Relativamente a estas vantagens, com maiores desenvolvimentos, Duarte Rodrigues Nunes, 2024, pp. 24 e ss.

Acresce que, no caso de o *CaaS* ser praticado por Estados, organizações terroristas ou outros grupos subversivos, os adquirentes das ferramentas e dos serviços ilícitos, ao levarem a cabo os crimes com utilização dessas ferramentas e/ou serviços, transformam-se - ainda que não deliberadamente - em agentes desses mesmos Estados, organizações terroristas ou grupos subversivos.

Todavia, além de ser um fator potenciador da danosidade/perigosidade de outras formas de criminalidade, o *CaaS* constitui igualmente uma ameaça *ex se*, sendo que, no caso das empresas (cada vez mais dependentes da informática[5]), o surgimento do *CaaS* veio intensificar a sua vulnerabilidade a ataques de um âmbito mais alargado de cibercriminosos, sobretudo no caso de micro, pequenas e médias empresas, que, em regra, não possuem sistemas de cibersegurança adequados[6]; o *CaaS* representa igualmente uma ameaça muito significativa à segurança nacional, já que agentes a soldo de Estados estrangeiros, organizações terroristas e/ou outros grupos subversivos podem usar esses serviços para realizar espionagem, sabotagem e outras operações cibernéticas contra infraestruturas críticas e sistemas governamentais (Ganguli, 2024, p. 2).

A fim de garantir tanto quanto possível a sua impunidade e a incolumidade da sua atividade e dos lucros obtidos, os agentes do *CaaS* utilizam igualmente as chamadas medidas antiforenses como as *passwords* "fortes", a encriptação de comunicações (especialmente, a encriptação ponta a ponta), a esteganografia, *firewalls*, *botnets*, VPNs, *proxies*, a *Dark Web,* programas como o TOR o *Freenet*, o I2P ou o *Mullvad*, o *Spoofing*, as criptomoedas, o *bulletproof hosting*, etc.[7]

Um outro aspeto a reter é a natureza frequentemente transnacional do *CaaS*, que dificulta grandemente a identificação dos agentes, a recolha de prova e a sua responsabilização (no mesmo sentido, Ganguli, 2024, p. 3), especialmente se os agentes estiverem em países (*v.g.*,

5 Fruto da evolução tecnológica, as empresas tiveram de passar a centrar muita da sua atividade económica no mundo virtual para, beneficiando da possibilidade de se dirigirem a um universo de consumidores muito mais vasto, aumentarem os seus lucros, para mais sabendo que, caso não o façam, perderão competitividade face às empresas que apostam na presença nesse mesmo mundo virtual.

6 Mas, mesmo no caso de grandes empresas e de grandes grupos económicos, que tenderão a possuir sistemas de cibersegurança muito avançados, existe sempre o perigo dos chamados ataques informáticos indiretos (*i.e.*, aqueles em que o cibercriminoso não ataca diretamente o sistema informático visado, antes explorando as vulnerabilidades de sistemas informáticos de terceiros que acedem ao sistema informático visado e que possuam sistemas de cibersegurança menos avançados).

7 Sobre as medidas antiforenses, Duarte Rodrigues Nunes, 2024, pp. 26 e ss.; Armando Dias Ramos, 2022, pp. 145 e ss.; David Silva Ramalho, 2017, pp. 152 e ss.; Europol, 2026, pp. 11 e ss.

a Rússia) cujas autoridades não colaborem com as suas congéneres dos países onde o processo corre os seus termos.

Finalmente, o uso de tecnologias baseadas na inteligência artificial (IA) veio tornar a engenharia social ainda mais eficaz (Europol, 2024, p. 8) e facilitar ainda mais a prática de crimes informáticos por parte de cibercriminosos com menores aptidões no plano da informática (Europol, 2026, p. 36), sendo que tais tecnologias são utilizadas pelos agentes do *CaaS*, na medida em que facilitam a produção de *malware* e incrementam a sua eficácia no cometimento de crimes.

Deste modo, o combate eficaz ao *CaaS* exige uma abordagem abrangente, que terá de incluir os meios proporcionados pela constante evolução da tecnologia, aperfeiçoamento da legislação e cooperação judiciária e policial internacional (Ganguli, 2024, p. 3). Essa abordagem não pode ser meramente reativa (*i.e.*, posterior à aquisição da *notitia criminis*), antes tendo de ser igualmente proativa (*i.e.*, prévia à obtenção da *notitia criminis*, sendo que uma das finalidades da investigação proativa é precisamente a obtenção de *notitiae criminum*)[8]. Uma das vertentes desta abordagem proativa é, precisamente, a compreensão da estrutura, do *modus operandi* e dos efeitos do *CaaS*, aspeto que é essencial para o desenvolvimento de contramedidas eficazes para proteger as pessoas, as empresas e as nações da crescente ameaça do cibercrime, altamente potenciada pelo surgimento do *CaaS* (Ganguli, 2024, p. 3).

II. A EVOLUÇÃO DO CIBERCRIME

No que tange à evolução do cibercrime, podemos dividi-la em três etapas, sendo que o *CaaS* corresponde à última dessas etapas[9].

A primeira etapa, que se iniciou ainda no século XX e que se prolongou até aos primeiros anos do século XXI, caracterizou-se essencialmente por condutas de acesso não autorizado a

8 Sobre a investigação proativa *vs.* investigação reativa e o *continuum* que deve existir entre ambas ao nível do combate ao crime organizado, ao terrorismo, ao crime económico e ao cibercrime, Duarte Rodrigues Nunes, 2019, pp. 255 e ss.

9 Relativamente à evolução do cibercrime, seguiremos de perto Ganguli, 2024, pp. 3 e ss., e Duarte Rodrigues Nunes, 2024, pp. 22 e ss.

sistemas informáticos e a dados neles armazenados e de vandalismo cibernético[10]. Os agentes do cibercrime atuavam, em regra, motivados pela curiosidade, afã de notoriedade ou desejo de causar prejuízos (seja por motivos ideológicos, religiosos, etc. ou pelo simples prazer de causar prejuízos aos visados), e não com a finalidade de obter lucro.

Na maioria das situações, as condutas consistiam em aceder, sem autorização, a sistemas informáticos e, eventualmente, aos dados neles armazenados ou acessíveis a partir deles, disseminação de vírus informáticos[11] e vandalização de páginas *web* (*defacing*). Os agentes atuavam individualmente ou em pequenos grupos "não organizados" e "não profissionalmente", independentemente da dimensão dos prejuízos e dos incómodos que as suas condutas causavam.

Nesta época histórica, a legislação de combate ao cibercrime era ainda muito rudimentar, sem prejuízo de a sua elaboração e aprovação serem um claro sinal da preocupação do legislador em responder, de forma o mais adequada possível, a um fenómeno criminoso ainda recente. No caso português, encontramos a Lei n.º 109/91, de 17 de agosto (Lei da criminalidade informática), que previa já alguns crimes informáticos nos seus arts. 4.º a 9.º (falsidade informática, dano relativo a dados ou programas informáticos, sabotagem informática, acesso ilegítimo, interceção ilegítima e reprodução ilegítima de programa protegido). Contudo, esta lei não continha quaisquer normas relativas à obtenção de prova digital (havendo que recorrer às normas do CPP que previam os meios de obtenção de prova "tradicionais" aplicáveis à obtenção de prova no mundo *offline*) nem relativas à cooperação internacional em matéria de combate ao cibercrime.

Todavia, ciente da evolução do cibercrime e do aumento da sua danosidade/perigosidade, bem como da existência de lacunas na legislação de diversos países em matéria de resposta ao cibercrime e de diferenças substanciais ao nível das ordens jurídicas de vários países igualmente no que tange ao combate ao cibercrime (o que dificulta de sobremaneira a cooperação judiciária internacional, que é essencial para uma resposta eficaz), o Conselho da Europa e a União Europeia (UE), com o objetivo de incrementar a eficácia da resposta ao Cibercrime, adotaram a Convenção sobre Cibercrime do Conselho da Europa, aberta à

10 Condutas subsumíveis aos crimes de acesso ilegítimo e de dano relativo a programas ou outros dados informáticos.

11 Em sentido amplo, incluindo vírus propriamente ditos, *worms*, etc.

assinatura em Budapeste em 23 de novembro de 2001 (CCiber)[12] (no caso do Conselho da Europa) e na Decisão-Quadro 2005/222/JAI, do Conselho, de 24 de fevereiro[13], relativa a ataques contra sistemas de informação, subsequentemente revogada e substituída pela Diretiva 2013/40/UE do Parlamento Europeu e do Conselho, de 12 de agosto de 2013, relativa a ataques contra os sistemas de informação e que substitui a Decisão-Quadro 2005/222/JAI do Conselho[14] (no caso da UE).

A segunda etapa, que se iniciou a partir do início do século XXI, caracterizou-se pelo surgimento de uma criminalidade informática essencialmente motivada pela obtenção de lucro, sem prejuízo de continuarem a existir situações em que os agentes atuavam movidos e "apenas" pela curiosidade, afã de notoriedade ou desejo de causar prejuízos.

O surgimento de uma criminalidade informática essencialmente motivada pela obtenção de lucro e usufruindo de novas oportunidades e nichos de mercado (*malware* cada vez mais sofisticado, *Dark Web*, programas altamente encriptados como o *Tor* ou o *Freenet*, criptomoedas, serviços de comunicações eletrónicas altamente encriptadas, etc.) trouxe várias novas realidades:

a) Cometimento de crimes informáticos visando a obtenção de lucro;
b) Profissionalização do cibercrime;
c) Envolvimento de organizações criminosas "tradicionais", organizações terroristas, criminosos de colarinho branco e Estados na prática de crimes informáticos;
d) Grupos organizados de cibercrime;
e) Mercados ilícitos *online*.

Relativamente ao cometimento de crimes informáticos visando a obtenção de lucro, com o incremento da atividade económica e financeira *online* e da utilização de sistemas de *home banking* e das compras *online*, aliado ao desenvolvimento tecnológico na área de informática (que proporcionou os meios necessários para a produção e disponibilização de *malware* cada vez mais sofisticado), surgiram novas atividades criminosas como o furto de identidade (para

12 E assinada por Portugal nessa mesma data, aprovada pela Resolução da Assembleia da República n.º 88/2009, de 15 de setembro, e ratificada pelo Decreto do Presidente da República n.º 91/2009, de 15 de setembro.
13 Publicada no JOUE n.º L 69, de 16/03/2005.
14 Publicada no JOUE n.º L 218, de 14/08/2013.

seguidamente levar a cabo fraudes utilizando a identidade de terceiros), o *Phishing* e outras formas de fraude *online* e o furto de dados confidenciais com a finalidade de obtenção de lucro através da sua ulterior cedência a terceiros mediante o pagamento de uma determinada quantia e, subsequentemente, o *Ransomware* e o *Cryptojacking*. Uma situação curiosa é a dos ataques *DDoS*, que, numa primeira fase, eram levados a cabo apenas com a finalidade de causar prejuízo ou, pelo menos, perturbar o mais possível a atividade do(s) visado(s), mas, a partir de uma determinada altura, passaram a ser levados a cabo também com o objetivo de extorquir dinheiro aos visados como condição para a cessação desses ataques (Duarte Rodrigues Nunes, 2024, p. 22; Europol, 2020, p. 32; Europol, 2021, p. 23).

No que diz respeito à profissionalização do cibercrime, fruto do cometimento de crimes informáticos visando a obtenção de lucro, surgiram cibercriminosos que passaram a fazer do cometimento de crimes informáticos modo de vida, retirando daí a totalidade ou uma parte substancial dos seus rendimentos, assistindo-se mesmo a casos de assaltantes de rua que, possuindo os conhecimentos informáticos necessários para tal, abandonaram a criminalidade *offline* e passaram a cometer crimes no mundo *online*, em que os proventos são tendencialmente superiores (ou muito superiores) e os riscos penais (em termos de probabilidades de punição e de severidade da punição) e de outra natureza (*v.g.*, os riscos decorrentes de uma reação da vítima ou de um terceiro) são sensivelmente menores.

Uma outra vertente da profissionalização do cibercrime foi o envolvimento de organizações criminosas "tradicionais", Estados, organizações terroristas e criminosos de colarinho branco na prática de crimes informático e o surgimento de grupos organizados de cibercrime, que, como veremos, possuem características diversas das organizações criminosas "tradicionais" e das organizações terroristas.

Começando pelo envolvimento de organizações criminosas "tradicionais"[15] (de que as máfias são o exemplo paradigmático, mas não o único), a utilização de meios informáticos permite uma melhor e mais eficiente comunicação entre a cúpula e as células da organização localizadas em locais diversos ou entre essas células (*maxime* no caso de organizações criminosas que possuam uma atuação transnacional), a realização - de forma menos "visível" - de operações de branqueamento com utilização de paraísos fiscais, a exploração nichos de

15 Sobre os conceitos de organização criminosa, associação criminosa, criminalidade económica organizada, "contiguidade" à organização criminosa e criminalidade organizada, Duarte Rodrigues Nunes, 2019, pp. 98, 131, 156, 187-188 e 191 e ss.

mercado (lícitos ou ilícitos) *online* ou mesmo *offline* de forma mais eficaz e uma proteção mais eficiente da organização, dos seus membros e do seu património face à atuação das autoridades (Duarte Rodrigues Nunes, 2021, pp. 29-30; Duarte Rodrigues Nunes, 2025a, pp. 45 e 47[16]; Armando Dias Ramos, 2022, p. 80; Joyce, 2026, *passim*; Gratteri/Nicaso, 2024, pp. 34 e ss., 62, 73-74 e 163 e ss.; Louise Shelley, 2003, pp. 304 e ss.; Europol, 2020, p. 51**).** No caso das organizações terroristas[17], o recurso às novas tecnologias de comunicação tanto pode servir para a proteção da organização como para a prossecução da sua finalidade terrorista (*v.g.,* realizar ataques terroristas, propaganda, captação de novos membros e de simpatizantes da causa terrorista, comunicação entre o núcleo central e as várias "células" independentes, coordenação entre os vários componentes da organização, obtenção de informações úteis para a organização, transferência de capitais, obtenção de lucro por via de fraudes cometidas através da Internet, etc.) **(**Duarte Rodrigues Nunes, 2021, p. 30; Dambruoso/Conti, 2023, pp. 229 e ss.; Cláudia Pina/José Vegar, 2025, pp. 81-82; Farber, 2025, *passim*; ONU, 2024, *passim*; Pierre Penalba/Abigaelle Penalba, 2022, pp. 105 e ss.; Michel Claise, 2024, pp. 72-73; Louise Shelley, 2003, pp. 304 e ss.; Sánchez Medero, 2012, pp. 74 e ss.), conferindo uma enorme rapidez e um anonimato absoluto ou quase absoluto às comunicações, potenciando a capacidade operacional da organização e dificultando enormemente a evitação de atentados terroristas, o desmantelamento de organizações terroristas, a perseguição e a punição os seus membros e apoiantes e o confisco do seu património de origem criminosa e/ou destinado a financiar a atividade terrorista e outras atividades conexas (Duarte Rodrigues Nunes, 2021, p. 30).

16 Mesmo no caso do Primeiro Comando da Capital (PCC), que é uma das organizações criminosas mais poderosas do Mundo e trafica droga à escala mundial, no caso das suas atividades licitas e ilícitas no território brasileiro, ainda que o Brasil seja o 2.º mercado interno de drogas no Mundo (apenas atrás dos EUA), o produto do tráfico de droga representa cerca de 5% das receitas anuais da criminalidade organizada (15.200 milhões de reais em 2022 e 2023), ao passo que o produto da venda de telemóveis furtados/roubados e das fraudes *online* representa cerca de 53% (148.129 milhões de reais em 2022 e 2023), estimando-se que, em 2023, o volume total das receitas do crime organizado no Brasil ascendeu a 348,1 mil milhões de reais (correspondentes a 56 mil milhões de euros) ["CRISE: COMO O PCC LAVOU BILHÕES NA FARIA LIMA", disponível em *https://www.youtube.com/watch?v=UtYgv3OMvLU* (visionado em 21/07/2026), e "O esquema bilionário do PCC descoberto pela PF", disponível em *https://www.youtube.com/watch?v=UiBrZPx3Ggk&t=11s* (visionado em 21/07/2026)].

17 Sobre os conceitos de terrorismo e de organização terrorista, Duarte Rodrigues Nunes, 2019, p. 170.

Ao nível da grande criminalidade económica[18] (embora seja uma realidade transversal, que inclui também a criminalidade organizada "tradicional" e o terrorismo), um dos domínios em que mais se lança mão dos meios informáticos para a proteção dos criminosos face às autoridades é o branqueamento de capitais, ao ponto de se afirmar que a informática é um meio *essencial* para o branqueamento e que o branqueamento só se consolidou como atividade conatural da criminalidade organizada com a possibilidade de recorrer às novas tecnologias e de se entender que existe uma relação de "conexão necessária" entre a criminalidade organizada, o branqueamento de capitais e a criminalidade informática[19] (Duarte Rodrigues Nunes, 2021, p. 31, Gutiérrez Francés, 2002**,** pp. 194-196 e 209, e Myriam Quéméner/Joël Ferry, 2009, p. 70).

No que tange aos grupos organizados de cibercrime, tais grupos criminosos (alguns dos quais atuam exclusivamente *online*, ao passo que outros combinam atividades *online* com atividades *offline*) apresentam um comportamento semelhante ao comportamento das organizações criminosas tradicionais e terroristas. Este aspeto verifica-se sobretudo no que tange à utilização de uma estrutura e de procedimentos concebidos para preservar o anonimato dos seus membros e evitar a atenção das autoridades através da implementação de medidas de segurança operacional para ocultar as suas identidades e atividades criminosas, medidas essas que são adaptadas ao *modus operandi* concreto das agências de aplicação da lei que correspondam ao tipo de serviços ilícitos que prestam e às especificidades desses mesmos serviços, sendo que, por exemplo, os *sites* de pornografia infantil costumam possuir medidas de segurança maiores do que os *sites* de venda de drogas ou de outros bens ilícitos (Duarte Rodrigues Nunes, 2024, pp. 39-40; Cláudia Pina/José Vegar, 2025, pp. 34 e ss. e 215 e ss.; ONU, 2022, pp. 16-17; Bussolati, 2023, pp. 261 e ss.).

Relativamente ao envolvimento de Estados no cibercrime, a atuação de entidades estatais (ou de cibercriminosos recrutados para esse fim) consiste, em regra, em atividades de ciberespionagem ou de ciberguerra (incluindo guerra económica e sabotagem), existindo igualmente casos em que a finalidade dos ciberataques é a obtenção de lucro para financiamento de determinados programas estatais, *maxime* programas de armamento (como

18 Sobre os conceitos de criminalidade económica e de criminalidade económica organizada, Duarte Rodrigues Nunes, 2019, pp. 143 e 152-153.

19 Falando-se, inclusivamente, em ciberbranqueamento, por oposição ao branqueamento "clássico" (Plantamura, 2023, pp. 941 e ss.).

sucede no caso da Coreia do Norte) [Duarte Rodrigues Nunes, 2024, pp. 24 e 35-36 (nota 53); Misha Glenny, 2012, pp. 212 e ss.; Cláudia Pina/José Vegar, 2025, pp. 79 e 173 e ss.; Holt *et. al.*, 2023, pp. 826 e ss.; Shloman, 2025; Ganguli, 2024, pp. 5-6 e 8-9; Pierre Penalba/Abigaelle Penalba, 2022, pp. 66-67).

Relacionados com todos os fatores que referimos, surgem os mercados ilícitos[20], que estão localizados na *Dark Web* - como sucede na maioria dos casos – ou em plataformas de serviços de mensagens instantâneas com elevado grau de criptografia. Cruzando os mercados ilícitos com os grupos organizados de cibercrime, os grupos que combinam atividades *online* com atividades *offline* levarão a cabo, por exemplo, atividades de tráfico de estupefacientes ou de armas e atos de violência física ou psíquica contra pessoas ou bens corpóreos, ao passo que os grupos que atuam exclusivamente *online* levarão a cabo, por exemplo, fraudes *online* ou ataques informáticos.

Em termos de Direito português, a entrada em vigor da Lei n.º 109/2009, de 15 de setembro, que transpôs para a nossa ordem jurídica a CCiber e a Decisão-Quadro n.º 2005/222/JAI e revogou a Lei n.º 109/2009, constitui uma resposta a esta segunda etapa da evolução do cibercrime. A Lei n.º 109/2009 manteve a punição dos mesmos tipos de crime que já constavam da Lei n.º 109/91, mas, dentro deles alargou o âmbito da punição das condutas que já eram punidas e acrescentou novas modalidade de condutas puníveis. Além disso, ao contrário da Lei n.º 109/91, a Lei n.º 109/2009 inclui normas relativas a meios de obtenção de prova específicos para a prova digital e normas relativas à cooperação internacional.

Finalmente, a terceira etapa consiste no surgimento do *CaaS*, enquanto modelo de negócio (ilícito) e que constituiu uma verdadeira revolução no plano do cibercrime, como veremos. Não é possível determinar com exatidão quando é que o *CaaS* surgiu, mas esta atividade criminosa terá as suas origens no surgimento de *kits* para campanhas de *Phishing* e de exploração de vulnerabilidades, sendo que os cibercriminosos que disponibilizavam esses *kits* perceberam rapidamente o potencial lucrativo dessa atividade criminosa, daí resultando a sua disseminação.

No que tange ao Direito português, a alteração da Lei n.º 109/2009 pela Lei n.º 79/2021, de 24 de novembro, que introduziu novos tipos de crime relacionados com a contrafação e o uso

20 Nos quais devemos incluir quer o fornecimento de bens/prestação de serviços de natureza ilícita quer o fornecimento de bens/prestação de serviços de natureza lícita, mas obtidos ou disponibilizados de forma ilícita.

abusivo de cartões e outros dispositivos de pagamento diversos do numerário, constitui uma resposta do legislador ao *CaaS*. Com a Lei n.º 79/2021, alargou-se o âmbito da punição de infrações que já constavam da Lei n.º 109/2009 e do Código Penal, sendo que algumas das condutas incriminadas são subsumíveis ao *CaaS* (sem prejuízo de, como veremos, a Lei n.º 109/2009 já punisse diversas condutas subsumíveis ao *CaaS*).

III. O *CYBERCRIME-AS-A-SERVICE*

Relativamente ao conceito, como referimos, o *CaaS* consiste na disponibilização de ferramentas e/ou serviços destinados à prática de crimes informáticos a qualquer pessoa disposta a pagar por eles.

Essas ferramentas e serviços incluem, por exemplo, *malware* (*kits* de *Ransomware*, *Phishing*, DDoS, *Cryptojacking*, *Infostealing*, *Exploit* e/ou *Sextortion, droppers*, furto de criptomoedsas, etc.), *Sim Boxes/SIM farms*, LLM malicioso[21], serviços de *Spoofing*, dados pessoais de terceiros, credenciais de acesso a sistemas e meios de pagamento (verdadeiros, mas obtidos através da prática de crimes), cartões e dispositivos de pagamento diversos do numerário contrafeitos, documentos contrafeitos, locação de *botnets*[22] para execução de ataques informáticos e *Cryptojacking*, dispositivos e serviços de comunicações com encriptação ponta a ponta (que permitem a realização de comunicações, por parte de criminosos, com baixos riscos de interceção pelas autoridades), prestação de serviços por *hackers* a soldo do "cliente", cursos *online* de atividades de cibercrime, disponibilização de serviços de *bulletptroof hosting*, VPNs e/ou *proxies*, plataformas ilícitas de *mixing* de criptomoedas, etc.[23] (Cláudia Pina/José Vegar, 2025, pp. 32, 39 e ss., 67-68, 74, 80, 124 e

21 *Large Language Models maliciosos*. O LLM malicioso é um modelo de IA destinado a permitir/facilitar a prática de crimes informáticos através da neutralização dos dispositivos de segurança do sistema informático atingido.

22 Com uma *botnet*, os cibercriminosos podem utilizar uma miríade de sistemas informáticos ao mesmo tempo para, por exemplo, levarem a cabo ataques informáticos (nomeadamente, ataques de DDoS, em que o ataque consiste em sobrecarregar o sistema informático com um número muito elevado de solicitações, de que resultará o seu entravamento ou, no mínimo, uma perturbação substancial do seu funcionamento) com maior eficácia e para conseguirem uma maior capacidade de mineração de criptomoedas, em ordem a receberem prémios (pagos em criptomoedas) por serem os primeiros a confirmar uma transação na *blockchain*.

23 Tendo em conta a miríade de ferramentas e serviços disponibilizados no âmbito do CaaS, existem modalidades específicas de CaaS consoante a ferramenta ou o serviço disponibilizado, como, por exemplo;

a) *Ransomware-as-a-Service* (RaaS);

b) *Phishing-as-a-Service* (PaaS);

127-128; Gratteri/Nicaso, 2024, pp. 158 e ss.; Ganguli, 2024, pp. 4-5 e 6-7; Tatiana Tropina, 2013, p 52; Pierre Penalba/Abigaelle Penalba, 2022, pp. 58 e ss. e 139 e ss.; Europol, 2020, pp. 6, 16-17, 22, 27, 57 e 59; Europol, 2021, pp. 9-10, 17-18 e 22-23; Europol, 2023a, pp. 7 e ss. e 13 e ss.; Europol, 2023b, pp. 6-7, 9-10, 11 e 13; Europol, 2024, pp. 8, 15-16, 18-19, 21-22 e 28; Europol, 2025, pp. 17-18 e 20 e ss.; Europol, 2026, pp. 8, 14-15, 19, 20, 21 e 23). Esta disponibilização consubstancia-se numa compra e venda, numa locação ou numa prestação de serviços que costumam ter lugar em mercados *online* ou fóruns sedeados na *Dark Web* ou em serviços de comunicações encriptadas, sendo os pagamentos realizados com criptomoedas (Ganguli, 2024, pp. 6, 7-8 e 9 e ss.; Cláudia Pina/José Vegar, 2025, pp. 38, 41, 67-68, 74, 80, 124 e 127-128; Europol, 2020, pp. 16-17, 22, 27 e 57; Europol, 2021, pp. 17-18; Europol, 2023a, pp. 10 e 13 e ss.; Europol, 2023b, pp. 6, 9-10, 11 e 13; Europol, 2024, pp. 8 e 15-16; Europol, 2025, p. 20).

Como referimos, o *CaaS* constitui uma nova etapa da evolução do cibercrime, permitindo que indivíduos que não possuem as capacidades necessárias para praticar crimes informáticos passem a dispor de meios que lhes permitem cometer tais crimes e incrementando a eficiência e a lucratividade da atividade criminosa desenvolvida pelos cibercriminosos tecnologicamente mais dotados e pelas organizações criminosas que disponibilizam as ferramentas e os serviços de *CaaS* (Europol, 2021, p. 17).

Na verdade, relativamente ao primeiro aspeto que referimos, podemos dizer que o *CaaS* "democratizou" o cibercrime, ao permitir que indivíduos com poucas aptidões informáticas levem a cabo, com enorme facilidade, atividades fraudulentas *online* (*v.g.*, *Phsishing, Ransomware, Cryptojacking*) e ataques informáticos (*v.g.*, ataques de DDoS ou de outra natureza contra infraestruturas críticas, instituições bancárias, etc.) em grande escala, com todas as consequências altamente negativas que tal pode acarretar em termos de segurança interna, intimidade/privacidade das pessoas, satisfação das necessidades essenciais dos cidadãos, estabilidade da economia, prejuízos económicos, etc.

---

c) DDoS-*as-a-Service* (DDoSaaS);
d) *Malware-as-a-Service* (MaaS);
e) *Droppers-as-a-Service*;
f) *Exploit-as-a-service*;
g) *Draining-as-a-Service*;
h) *Mixing-as-a-Service*.

E, compreensivelmente, essa "democratização" levou a um aumento substancial do número de cibercriminosos e do número de crimes informáticos cometidos, dificultando sensivelmente a prevenção e a repressão do cibercrime.

Relativamente ao segundo aspeto, o *CaaS* proporciona aos que disponibilizam as ferramentas e os serviços ilícitos lucros avultados, que serão depois usufruídos pelo(s) agente(s) do crime (seja uma pessoa física, um ente coletivo, uma organização criminosa ou um Estado) e/ou reinvestidos em novas atividades criminosas e/ou utilizados para corromper agentes públicos ou para pagar a colaboradores externos de organizações criminosas; no caso de Estados (*v.g.*, a Coreia do Norte), o lucro é utilizado para financiar programas de armamento e, no caso de organizações terroristas e de outros grupos subversivos, o lucro será utilizado para financiar a atividade terrorista ou subversiva.

Ainda no caso de o *CaaS* ser praticado por Estados, organizações terroristas ou outros grupos subversivos, os adquirentes das ferramentas e dos serviços ilícitos, ao levarem a cabo os crimes com utilização dessas ferramentas e/ou serviços, transformam-se, ainda que não deliberadamente, em agentes desses mesmos Estados, organizações terroristas ou grupos subversivos. Na verdade, ao levarem a cabo ataques informáticos contra infraestruturas críticas, o sistema económico-financeiro ou mesmo contra um grande número de cidadãos, os cibercriminosos acabam por causar prejuízos similares aos danos que os ataques informáticos que são habitualmente levados a cabo por Estados, organizações terroristas ou grupos subversivos costumam causar. Por isso, no caso de Estados, organizações terroristas e outros grupos subversivos, o *CaaS*, além de proporcionar lucros muito substanciais, também "fornece" novos "soldados" a essas entidades.

No que concerne aos agentes do *CaaS*, temos os seguintes:

a) Criadores (*developers*): criam as ferramentas maliciosas e/ou prestam os serviços de assistência/"formação" de cibercriminosos;

b) Provedores de acesso [*initial access brokers* (IABs)]: proporcionam/facilitam o acesso não autorizado a sistemas informáticos, redes ou contas *online* através da disponibilização de credenciais "furtadas" ou de outras formas de acesso ilegítimo; e

c) Operadores (*operators*): executam os ataques informáticos ou as fraudes *online* utilizando as ferramentas adquiridas a terceiros.

Além disso, cumpre referir, ainda, aqueles que, conscientemente, auxiliam as atividades de *CaaS* de terceiros através da disponibilização de plataformas *online* e outros serviços que facilitam as atividades de *CaaS* (serviços de mensagens encriptadas, serviços de *bulletproof hosting*, VPNs e/ou *proxies*, corretagem de criptomoedas, etc.) (Europol, 2020, pp. 22; Europol. 2021, pp. 8 e 18; Europol. 2023b, pp. 16-17).

Em muitos casos, tais serviços são disponibilizados com a finalidade de permitir a sua utilização para fins criminosos. Mas também existem casos em que, apesar de não ser essa a sua finalidade, aqueles que disponibilizam esses serviços não podem ignorar que, pelas suas características, serão utilizados, com elevada probabilidade, para fins criminosos e, ainda assim, não adotam quaisquer medidas destinadas a prevenir a utilização dos seus serviços para fins criminosos e, quando solicitada a sua cooperação pelas autoridades, recusam-na ou dificultam-na ao máximo, com o pretexto de estarem obrigados a proteger a privacidade dos seus clientes (Europol, 20220, p. 22 e 57; Europol. 2021, p. 18).

IV. A AMEAÇA DO *CYBERCRIME-AS-A-SERVICE*

Como referimos, o *CaaS* revolucionou o cibercrime, que deixou de ser um exclusivo de *hackers* altamente capacitados, passando a prática de crimes informáticos (incluindo ataques informáticos em grande escala) a estar ao alcance de criminosos que, sem a disponibilização de tais ferramentas e/ou serviços, não teriam capacidade para tal.

Deste modo, além de o número de agentes ou potenciais agentes de crimes informáticos ter aumentado substancialmente, o número de crimes informáticos tem registado um forte aumento[24]; e o mesmo vale para a sofisticação e a frequência dos ataques informáticos.

De facto, o universo dos alvos dos crimes informáticos (com especial destaque para o *Ransmoware* - incluindo contra infraestruturas críticas - o *Phishing* e outras formas de fraude *online*, o "furto" de dados pessoais ou de credenciais de acesso a sistemas e meios de pagamento, a espionagem e os ataques DDoS contra infraestruturas críticas) cuja prática é possibilitada pela disponibilização de ferramentas e/ou serviços através de plataformas de

24 No caso de Portugal, o número de denúncias tem aumentado substancialmente: 193 denúncias em 2019, 544 denúncias em 2020, 1160 denúncias em 2021, 2124 denúncias em 2022, 2916 denúncias em 2023, 3973 denúncias em 2024 e 4497 denúncias em 2025 (Gabinete Cibercrime da Procuradoria Geral da República, 2026, p. 5).

CaaS é vasto, incluindo pessoas singulares (muitas vezes, em grande número), empresas (PMEs, bancos, grandes empresas ou mesmo multinacionais) e outros organismos públicos ou privados.

Além disso, os prejuízos causados não se limitam a danos patrimoniais ou ao "furto" de propriedade intelectual, incluindo também a exposição de dados sensíveis e a interrupção de serviços críticos (saúde, abastecimento de água, eletricidade ou gás, segurança aérea e/ou rodoviária, sistema de Justiça e/ou de segurança interna, etc.), de que poderão resultar mortes, perigo para a vida ou ofensas graves à integridade física, além de atentados graves à intimidade/privacidade e à segurança externa e interna dos Estados (Ganguli, 2024, pp. 2, 12 e 19 e ss.).

Se é verdade que a ameaça do cibercrime em geral foi altamente incrementada pelo *CaaS*, a ameaça, quer do *CaaS* quer - consequentemente - do cibercrime em geral, tem sido igualmente incrementada de forma substancial pelo uso de IA, sendo que as ferramentas de IA generativa como o *FraudGPT*, o *BlackmailerV3* e o *ElevenLabs* são cada vez mais utilizadas para criar *malware* e para aperfeiçoar a engenharia social (tornando-a mais credível para as vítimas), acelerando e ocultando esquemas de fraude *online* (Europol, 2026, pp. 15, 23, 24 e 37).

V. OS DESAFIOS QUE O *CYBERCRIME-AS-A-SERVICE* COLOCA ÀS AUTORIDADES

Como referimos, os agentes do *CaaS* utilizam as chamadas medidas antiforenses como as *passwords* "fortes", a encriptação de comunicações (especialmente, a encriptação ponta a ponta), a esteganografia, *firewalls*, *botnets*, VPNs, *proxies*, a *Dark Web,* programas como o TOR o *Freenet*, o I2P ou o *Mullvad*, o *Spoofing*, as criptomoedas, o *bulletproof hosting*, etc., cuja finalidade é dificultar a deteção da prática de crimes informáticos pelas autoridades (sabendo igualmente que muitos crimes informáticos não são denunciados pelas vítimas), a identificação dos agentes dos crimes e a investigação criminal e patrimonial.

Além da utilização de *firewalls*, *botnets*, VPNs e/ou *proxies* para dissimular a sua identificação e localização, bem como dificultar a recolha de provas, os agentes do *CaaS*, com as mesmas finalidades, comercializam as suas ferramentas e serviços em locais

especialmente protegidos contra a atuação das autoridades, como os *marketplaces* e os fóruns da *Darkweb*[25], os serviços de comunicações encriptadas ponta a ponta e as plataformas de *bulletproof hosting* (Cláudia Pina/José Vegar, 2025, p. 124; Ganguli, 2024, pp. 2, 7-8 e 9 e ss.; Europol, 2020, p. 58; Europol, 2021, pp. 8 e 18; Europol, 2023a, pp. 16-17). De facto, à sua enorme racionalidade em termos de cálculo de custos *vs.* benefícios (*i.e.*, as vantagens económicas e/ou de outra natureza que poderão obter através da prática de crimes *vs.* os riscos penais em que incorrem), os cibercriminosos juntam especiais cautelas e um *modus operandi* que visa, por um lado, aumentar a sua eficácia criminosa e, por outro, afastar ou diminuir drasticamente as probabilidades de punição.

Um outro aspeto do *CaaS* que dificulta de sobremaneira a sua prevenção e a sua repressão é a utilização de criptomoedas como meio de pagamento das ferramentas e serviços, uma vez que a descentralização e o anonimato das operações com criptomoedas dificultam de sobremaneira a monitorização dessas operações, a recolha de prova e a localização e o congelamento dos respetivos ativos pelas autoridades (Cláudia Pina/José Vegar, 2025, pp. 106 e 124; Ganguli, 2024, pp. 2 e 12 e ss.; Europol, 2020, p. 17; Europol, 2023b, p. 13**;** Europol, 2026, pp. 11 e 14-15**).**

Também o carácter tendencialmente transnacional do *CaaS* dificulta a resposta ao *CaaS* (Ganguli, 2024, p. 23). Na verdade, não poucas vezes, as plataformas de *CaaS* estão localizadas num país, ao passo que as ferramentas e os serviços nelas disponibilizados são utilizados para atingir vítimas em todo o Mundo, sendo que os cibercriminosos poderão estar localizados em qualquer país[26] (Europol, 2023a, p. 17). As dificuldades resultam da necessidade de identificar previamente os países em causa, solicitar a cooperação entre as autoridades judiciárias e policiais dos vários países envolvidos (sendo que existem Estados que não cooperam ou que não cooperam de forma expedita) e das diferenças entre as legislações internas dos vários Estados em causa no que tange à proteção de dados e da privacidade e aos meios de investigação (Europol, 2021, pp. 8 e 18).

25 Os mercados da *Dark Web* estão cada vez mais fragmentados, havendo uma clara preferência dos cibercriminosos por plataformas mais especializadas em atividades criminosas específicas, sendo que, apesar das constantes operações de desmantelamento e dos esforços das autoridades, esses mercados demonstram uma resiliência notável, surgindo rapidamente novos mercados para substituir os que foram desmantelados; por sua vez, os fóruns da *Dark Web* funcionam como centros de migração para novos mercados da *Dark Web* após o encerramento de um mercado importante pelas autoridades (Europol, 2021, p. 36; Europol, 2026, pp. 11 e ss.).
26 Inclusivamente num país diverso daquele onde está localizada a plataforma e daquele(s) em que as ferramentas e serviços são utilizadas.

Por todos estes fatores, os agentes do *CaaS* podem levar a cabo uma atividade criminosa altamente rendível e com riscos penais diminutos (Ganguli, 2024, p. 2), pois a deteção das atividades de *CaaS*, a identificação dos seus agentes, a recolha de prova e a localização e o congelamento dos ativos relacionados com o crime são extremamente difíceis.

VI. A PUNIÇÃO DO *CYBERCRIME-AS-A-SERVICE* NA LEI PENAL PORTUGUESA

Tendo em conta os exemplos de condutas subsumíveis ao *CaaS*, encontramos na Lei n.º 109/2009, de 15 de setembro, diversas incriminações que incluem condutas subsumíveis ao CaaS.

Assim, desde logo, as condutas p. e p. pelos arts. 3.º, n.º 4, 4.º, n.º 3, 5.º, n.º 2, 6.º, n.ºs 2 e 3, e 7.º, n.º 3, da Lei n.º 109/2009, em que o legislador pune condutas que consistem na produção, aquisição, comercialização ou disponibilização por qualquer outra forma, disseminação ou detenção de dispositivos[27] destinados a (1) interferir num tratamento informático de dados, com vista à produção de dados ou documentos não genuínos, (2) causar danos em dados informáticos alheios, (3) entravar o funcionamento de sistemas informáticos alheios e/ou (4 e 5) aceder a sistemas informáticos e/ou a comunicações informáticas alheios.

Do mesmo modo, com a Lei n.º 79/2021, de 24 de novembro, o legislador agravou e alargou o âmbito da produção, aquisição, comercialização ou disponibilização por qualquer outra forma, disseminação ou detenção de cartões ou outros dispositivos de pagamento diversos do numerário contrafeitos e de dispositivos destinados à contrafação de cartões ou outros dispositivos de pagamento diversos do numerário (cfr. arts. 3.º-C e 3.º-D da Lei n.º 109/2009)[28].

E, por fim, ainda por via da introdução do art. 3.º-E na Lei n.º 109/2009[29], através da Lei n.º 79/2021, o legislador passou a punir a aquisição e a comercialização ou disponibilização por qualquer outra forma de cartões ou outros dispositivos de pagamento diversos do numerário

27 De *hardware* ou *software* (Duarte Rodrigues Nunes, 2024, p. 88).
28 Sobre a conexão entre estes tipos de crime e o CaaS, Duarte Rodrigues Nunes, 2024, pp. 168 e 188.
29 Relativamente à conexão entre este tipo de crime e o CaaS, Duarte Rodrigues Nunes, 2024, p. 207.

verdadeiros, mas obtidos por via da prática de crimes de dano relativo a programas ou outros dados informáticos, sabotagem informática, acesso ilegítimo ou interceção ilegítima[30].

Contudo, a nossa lei apresenta algumas lacunas e matéria de combate ao *CaaS*. Em primeiro lugar, também deveriam ser previstas modalidades de conduta típica análogas às previstas nos arts. 3.º, n.º 4, 4.º, n.º 3, 5.º, n.º 2, 6.º, n.ºs 2 e 3, e 7.º, n.º 3, da Lei n.º 109/2009 noutros tipos de crime, como, por exemplo, os crimes de burla informática (p. e p. pelo art. 221.º, n.º 1, do CP), extorsão[31] (p. e p. pelo art. 223.º do CP) e abuso de cartão de garantia ou de cartão, dispositivo ou dados de pagamento (nas modalidades de conduta típica previstas nas als. c) e d) do n.º 1 do art. 225.º do CP).

Em segundo lugar, deveria ser prevista a punição, enquanto incriminação autónoma (caso contrário, tais agentes, quando muito, apenas poderão ser punidos a título de cumplicidade) da conduta daqueles que, conscientemente, auxiliam as atividades de *CaaS* de terceiros através da disponibilização de plataformas *online* e outros serviços que facilitam as atividades de *CaaS* (serviços de mensagens encriptadas, serviços de *bulletproof hosting*, VPNs e/ou *proxies*, corretagem de criptomoedas, etc.), pois, como referimos, em muitos casos, tais serviços são disponibilizados com a finalidade de permitir a sua utilização para fins criminosos e, noutros, apesar de não ser essa a sua finalidade, aqueles que disponibilizam esses serviços não podem ignorar que, pelas suas características, serão utilizados, com elevada probabilidade, para fins criminosos e, ainda assim, não adotam quaisquer medidas destinadas a prevenir a utilização dos seus serviços para fins criminosos (designadamente, procedimentos de *Know Your Customer*) e, quando solicitada a sua cooperação pelas autoridades, recusam-na ou dificultam-na ao máximo, com o pretexto de estarem obrigados a proteger a privacidade dos seus clientes.

E, em terceiro lugar, tendo em conta que o *CaaS* é uma atividade criminosa suscetível de gerar proventos muito substanciais para os seus agentes e que a investigação do *CaaS* é extremamente difícil (inclusive quanto ao estabelecimento de um nexo entre o património do agente e a prática de crimes concretos), todos os tipos de crime subsumíveis ao *CaaS* deveriam constar do catálogo de crimes do art. 1.º, n.º 1, da Lei n.º 5/2002, de 11 de janeiro,

30 Para uma crítica à limitação do catálogo de crimes, Duarte Rodrigues Nunes, 2024, pp. 214-215, e Pedro Dias Venâncio, 2022, p. 109.

31 Um dos tipos de crime praticados pelos agentes do *Ransmoware* e dos ataques de DDoS (quando a sua finalidade é a obtenção do pagamento de um resgate).

em ordem a permitir o confisco do valor incongruente (anteriormente designado por "confisco alargado") nos termos do art. 7.º e ss. desse diploma[32] e o confisco alargado nos termos dos arts. 6.º-A ou 6.º-B, igualmente da Lei n.º 5/2002[33].

VII. CONCLUSÕES

A) O *CaaS* consiste na disponibilização de ferramentas e/ou serviços destinados à prática de crimes informáticos a qualquer pessoa disposta a pagar por eles.

B) Essas ferramentas e serviços incluem, por exemplo, *malware* (*kits* de *Ransomware*, *Phishing*, DDoS, *Cryptojacking*, *Infostealing*, *Exploit* e/ou *Sextortion, droppers*, furto de criptomoedsas, etc.), *Sim Boxes/SIM farms*, *malicious* LLM, serviços de *Spoofing*, dados pessoais de terceiros, credenciais de acesso a sistemas e meios de pagamento, cartões e dispositivos de pagamento diversos do numerário contrafeitos, documentos contrafeitos, locação de *botnets* para execução de ataques informáticos e *Cryptojacking*, dispositivos e serviços de comunicações com encriptação ponta a ponta (que permitem a realização de comunicações, por parte de criminosos, com elevadas garantias de não interceção pelas autoridades), prestação de serviços por *hackers* a soldo do "cliente", cursos *online* de atividades de cibercrime, disponibilização de serviços de *bulletptroof hosting*, VPNs e/ou *proxies*, plataformas ilícitas de mistura de criptomoedas, etc.

C) O *CaaS* revolucionou o cibercrime ao proporcionar a indivíduos com menores aptidões em matéria de informática meios que lhes permitem cometer crimes informáticos que, de outro modo, não conseguiriam praticar, levando a um aumento substancial do número de cibercriminosos e do número de crimes informáticos cometidos, dificultando sensivelmente a prevenção e a repressão

32 Sobre a nossa crítica à excessiva limitação do catálogo do confisco do valor incongruente, com maiores desenvolvimentos, Duarte Rodrigues Nunes, 2025b, pp. 15-16, em que é também referido o CaaS.
33 Para uma crítica à previsão na Lei n.º 5/2002 (e não no Código Penal), à existência de um catálogo de crimes e ao facto de o catálogo do confisco alargado ser diverso do catálogo do confisco do valor incongruente, Duarte Rodrigues Nunes, 2026, pp. 42-43 e 46.

do cibercrime e de outras formas de criminalidade em que sejam usados meios informáticos para cometer crimes.

D) O *CaaS* tornou-se num modelo de negócio similar ao negócio legítimo do *Software-as-a-Service*, dele se distinguido pelo facto de as ferramentas e/ou os serviços disponibilizados no âmbito do *CaaS* serem destinados a permitir/facilitar a ulterior prática de crimes, *i..e.*, maliciosos.

E) À semelhança do que sucede com outros mercados ilícitos *online*, o *CaaS* tira partido das características dos meios informáticos no que tange à dificultação da deteção dos crimes e da investigação criminal e à facilitação da preparação e execução dos crimes e do apagamento dos seus vestígios, constituindo uma atividade criminosa que acarreta riscos penais muito baixos para os criminosos.

F) Os agentes do *CaaS* utilizam as chamadas medidas antiforenses como as *passwords* "fortes", a encriptação de comunicações (especialmente, a encriptação ponta a ponta), a esteganografia, *firewalls*, *botnets*, VPNs, *proxies*, a *Dark Web,* programas como o TOR o *Freenet*, o I2P ou o *Mullvad*, o *Spoofing*, as criptomoedas, o *bulletproof hosting*, etc., cuja finalidade é dificultar a deteção da prática de crimes informáticos pelas autoridades (sabendo igualmente que muitos crimes informáticos não são denunciados pelas vítimas), a identificação dos agentes dos crimes e a investigação criminal e patrimonial.

G) Além de acarretar riscos penais muito pouco substanciais, o *CaaS* é uma atividade criminosa extremamente lucrativa, sendo os lucros obtidos usufruídos pelo agente do crime (que pode ser uma ou mais pessoas físicas agindo de uma forma "não organizada", um ente coletivo, uma organização criminosa ou terrorista ou um Estado), reinvestidos em novas atividades criminosas e/ou utilizados para corromper agentes públicos ou para pagar a colaboradores externos de organizações criminosas; no caso de Estados (*v.g.*, a Coreia do Norte), o lucro é utilizado para financiar programas de armamento ou operações ilícitas e, no caso de organizações terroristas e de outros grupos subversivos, o lucro será utilizado para financiar a atividade terrorista ou subversiva.

H) Os adquirentes das ferramentas e dos serviços ilícitos, ao levarem a cabo os crimes com utilização dessas ferramentas e/ou serviços, transformam-se – ainda

que não deliberadamente - em agentes desses mesmos Estados, organizações terroristas ou grupos subversivos.

I) O combate eficaz ao *CaaS* exige uma abordagem abrangente, que terá de incluir os meios proporcionados pela constante evolução da tecnologia, aperfeiçoamento da legislação e cooperação judiciária e policial internacional.

J) Essa abordagem não pode ser meramente reativa (*i.e.*, posterior à aquisição da *notitia criminis*), antes tendo de ser igualmente proativa (*i.e.*, prévia à obtenção da *notitia criminis*, sendo que uma das finalidades da investigação proativa é precisamente a obtenção de *notitiae criminum*).

K) No que concerne aos agentes do *CaaS*, existem os criadores (*developers*), os provedores de acesso [*initial access brokers* (IABs)], os operadores (*operators*) e aqueles que, conscientemente, auxiliam as atividades de *CaaS* de terceiros através da disponibilização de plataformas *online* e outros serviços que facilitam as atividades de *CaaS* (serviços de mensagens encriptadas, serviços de *bulletproof hosting*, VPNs e/ou *proxies*, corretagem de criptomoedas, etc.).

L) Em muitos casos, tais serviços são disponibilizados com a finalidade de permitir a sua utilização para fins criminosos, mas existem igualmente casos em que, apesar de não ser essa a sua finalidade, aqueles que disponibilizam esses serviços não podem ignorar que, pelas suas características, serão utilizados, com elevada probabilidade, para fins criminosos e, ainda assim, não adotam quaisquer medidas destinadas a prevenir a utilização dos seus serviços para fins criminosos e, quando solicitada a sua cooperação pelas autoridades, recusam-na ou dificultam-na ao máximo, com o pretexto de estarem obrigados a proteger a privacidade dos seus clientes.

M) Os prejuízos causados pelo *CaaS* incluem danos patrimoniais, o "furto" de propriedade intelectual, a exposição de dados sensíveis e a interrupção de serviços críticos (saúde, abastecimento de água, eletricidade ou gás, segurança aérea e/ou rodoviária, sistema de Justiça e/ou de segurança interna, etc.), de que poderão resultar mortes, perigo para a vida ou ofensas graves à integridade física, além de atentados graves à intimidade/privacidade e à segurança externa e interna dos Estados.

N) A ameaça do *CaaS* tem sido igualmente incrementada de forma substancial pelo uso de IA, sendo que as ferramentas de IA generativa como o *FraudGPT*, o *BlackmailerV3* e o *ElevenLabs* são cada vez mais utilizadas para criar *malware* e para aperfeiçoar a engenharia social (tornando-a mais credível para as vítimas), acelerando e ocultando esquemas de fraude *online*.

O) A Lei portuguesa inclui a punição de condutas subsumíveis ao *CaaS* nos arts. 3.º, n.º 4, 3.º-C, 3.º-D, 3.º-E, 4.º, n.º 3, 5.º, n.º 2, 6.º, n.ºs 2 e 3, e 7.º, n.º 3, da Lei n.º 109/2009, mas apresenta algumas lacunas e matéria de combate ao *CaaS*.

P) Deveriam ser previstas, na nossa Lei, modalidades de conduta típica análogas às previstas nos arts. 3.º, n.º 4, 4.º, n.º 3, 5.º, n.º 2, 6.º, n.ºs 2 e 3, e 7.º, n.º 3, da Lei n.º 109/2009 noutros tipos de crime, como, por exemplo, os crimes de burla informática, extorsão e abuso de cartão de garantia ou de cartão, dispositivo ou dados de pagamento (nas modalidades de conduta típica previstos nas als. c) e d) do n.º 1 do art. 225.º do CP).

Q) Deveria ser igualmente prevista a punição, enquanto incriminação autónoma (caso contrário, tais agentes, quando muito, apenas poderão ser punidos a título de cumplicidade) das condutas de auxílio doloso (incluindo com dolo eventual) às atividades de *CaaS* de terceiros através da disponibilização de plataformas *online* e outros serviços.

R) Por fim, todos os tipos de crime subsumíveis ao *CaaS* previstos na nossa ordem jurídica deveriam constar do catálogo de crimes do art. 1.º, n.º 1, da Lei n.º 5/2002, de 11 de janeiro, em ordem a permitir o confisco do valor incongruente (anteriormente designado por "confisco alargado") nos termos do art. 7.º e ss. desse diploma e o confisco alargado nos termos dos arts. 6.º-A ou 6.º-B, igualmente da Lei n.º 5/2002.

BIBLIOGRAFIA

BUSSOLATI, Nicolò (2023) - "*L' associazione per delinquere "informatica"*", *in* Cybercrime, II Edição, pp. 261 e ss., UTET Wolters Kluwer, Milão, 2023.

CLAISE, Michel (2024) - Combattre la criminalité : une urgence démocratique, Éditions Racine, Bruxelas, 2024.

DAMBRUOSO, Stefano/Conti, Francesco (2023) - "*Delitti con finalità di terrorismo con specifiche affravanti "technologiche"*", *in* Cybercrime, II Edição, pp. 229 e ss., UTET Wolters Kluwer, Milão, 2023.

EUROPOL (2020) - Internet Organised Crime Threat Assessment 2020, *in https://www.europol.europa.eu/publication-events/main-reports/internet-organised-crime-threat-assessment-iocta-2020#downloads* (acedido em 16/07/2026).

EUROPOL (2021) - Internet Organised Crime Threat Assessment 2021, *in https://www.europol.europa.eu/publication-events/main-reports/internet-organised-crime-threat-assessment-iocta-2021#downloads* (acedido em 16/07/2026).

EUROPOL (2023a) - Cyber-Attacks: The apex of crime-as-a-service, *in https://www.europol.europa.eu/publication-events/main-reports/internet-organised-crime-threat-assessment-iocta-2021#downloads* (acedido em 16/07/2026).

EUROPOL (2023b) - Internet Organised Crime Threat Assessment 2023, *in https://www.europol.europa.eu/publication-events/main-reports/internet-organised-crime-threat-assessment-iocta-2023#downloads* (acedido em 16/07/2026).

EUROPOL (2024) - Internet Organised Crime Threat Assessment 2024, *in https://www.europol.europa.eu/publication-events/main-reports/internet-organised-crime-threat-assessment-iocta-2024#downloads* (acedido em 16/07/2026).

EUROPOL (2025) - Internet Organised Crime Threat Assessment 2025, *in https://www.europol.europa.eu/publication-events/main-reports/internet-organised-crime-threat-assessment-iocta-2025#downloads* (acedido em 16/07/2026).

EUROPOL (2026) - Internet Organised Crime Threat Assessment 2026, *in https://www.europol.europa.eu/publication-events/main-reports/internet-organised-crime-threat-assessment-iocta-2026#downloads* (acedido em 16/07/2026).

FARBER, Shai (2025) - "*The evolving nexus of cybercrime and terrorism: A systematic review of convergence and policy implications*", *in* Security Journal, 38 (2025), pp. 28 e ss., *in https://link.springer.com/article/10.1057/s41284-025-00471-7* (acedido em 16/07/2026).

GABINETE CIBERCRIME DA PROCURADORIA GERAL DA REPÚBLICA (2026) - Nota Informativa CIBERCRIME: DENÚNCIAS RECEBIDAS 2025, *in https://cibercrime.ministeriopublico.pt/sites/default/files/2026-02/2026.02.11-denuncias-de-cibercrime-2025.pdf* (acedido em 20/07/2026).

GANGULI, Prithwish (2024) - The Rise of Cybercrime-as-a-Service: Implications and Countermersures, *in https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4959188* (acedido em 17/02/2026).

GLENNY, Misha (2012) - Darkmarket, Como os Hackers se tornaram a nova Máfia, Civilização, Lisboa, 2012.

GRATTERI, Nicola/Nicaso, Antonio (2024) - Una Cosa Sola, Mondadori, Roma, 2024.

GUTIÉRREZ FRANCÉS, Mariluz (2002) - "*Las altas tecnologías de la información al servicio del blanqueo de capitales transnacional*", *in* Blanqueo de Dinero y Corrupción en el Sistema Bancario, Delitos Financieros, Fraude y Corrupción en Europa, Volume II, pp. 193 e ss., Ediciones Universidad de Salamanca, Salamanca, 2002.

HOLT, Thomas J./Griffith, Mae/Turner, Noah/Greene-Colozzi, Emily/Chermak, Steven/Freilich, Joshua D. (2023) - "*Assessing nation-state-sponsored cyberattacks using aspects of Situational Crime Prevention*", *in* Criminology & Public Policy, 22, pp. 825 e ss., *in https://onlinelibrary.wiley.com/doi/full/10.1111/1745-9133.12646* (acedido em 17/06/2026).

JOYCE, Thomas (2026) - "*Cybercrime and Organised Crime: Distinct Criminal Forms in the Digital Era*", *in* Criminology - Understanding Crime, Criminality, and Justice in a Changing World, pp. 1 e ss., *in https://www.intechopen.com/chapters/1248970#* (acedido em 17/06/2026).

NUNES, Duarte Rodrigues (2019) - O problema da admissibilidade dos métodos "ocultos" de investigação criminal como instrumento de resposta à criminalidade organizada, Gestlegal, Coimbra, 2019.

NUNES, Duarte Rodrigues (2021) - Os meios de obtenção de prova previstos na Lei do Cibercrime, 2.ª Edição, Gestlegal, Coimbra, 2021.

NUNES, Duarte Rodrigues (2024) - Os crimes previstos na Lei do Cibercrime e a responsabilidade dos entes coletivos, 2.ª Edição, Gestlegal, Coimbra, 2024.

NUNES, Duarte Rodrigues (2025a) - "*O Primeiro Comando da Capital (PCC). História, estrutura, modus operandi, e atividades criminosas. O desafio para as autoridades portuguesas*", *in* Revista do Ministério Público, n.º 184 e ss., pp. 9 e ss., 2025.

NUNES, Duarte Rodrigues (2025b) - Lei de Combate à Criminalidade Organizada e Económico-Financeira Comentada e Anotada com a Jurisprudência, Gestlegal, Coimbra, 2025.

NUNES, Duarte Rodrigues (2026) - "*Pronúncia em sede de Consulta Pública (Proposta de Lei n.º 50/XVII/1)*", *in* Julgar Digital, julho de 2026, *in https://julgar.pt/pronuncia-em-sede-de-consulta-publica-proposta-de-lei-n-o-50-xvii-1/* (acedido em 19/06/2026).

ORGANIZAÇÃO DAS NAÇÕES UNIDAS (2022) - Compendio de Ciberdelincuencia Organizada, 2.ª Edição, Viena, 2022, *in https://www.unodc.org/documents/organized-crime/tools_and_publications/21-05345S_eBook.pdf* (acedido em 17/06/2026).

ORGANIZAÇÃO DAS NAÇÕES UNIDAS (2024) - Beneath the Surface: Terrorist and Violent Extremist use of the Dark Web and Cybercrime-as-a-Service for Cyber-Attacks, 2024, *in https://www.un.org/counterterrorism/sites/default/files/dw_beneath_the_surface_update.pdf* (acedido em 17/06/2026).

PENALBA, Pierre/Penalba, Abigaelle (2022) - Darknet, la voyage qui fait peur, Albin Michel, Paris, 2022.

PINA, Cláudia/Vegar, José (2025) - Cibercrime, Desassossego, Lisboa, 2025.

PLANTAMURA, Vit (2023) - "*Il cibericiclaggio*", *in* Cybercrime, II Edição, pp. 925 e ss., UTET Wolters Kluwer, Milão, 2023.

QUÉMÉNER, Myriam/Ferry, Joël (2009) - Cybercriminalité, Défi mondial, 2.ª Edição, Economica, Paris, 2009.

RAMALHO, David Silva (2017) - Métodos Ocultos de Investigação Criminal em Ambiente Digital, Almedina, Coimbra, 2017.

RAMOS, Armando Dias (2022) - O Agente Encoberto Digital, Meios especiais e técnicos de investigação criminal, Almedina, Coimbra, 2022.

SÁNCHEZ MEDERO, Gema (2012) - "*Ciberespacio y el crimen organizado. Los nuevos desafíos del siglo XXI*", *in* Revista Enfoques, Volume X, n.º 16, pp. 71 e ss., 2012.

SHELLEY, Louise (2003) - "*Organized Crime, Terrorism and Cybercrime*", *in* Security Sector Reform: Institutions, Society and Good Governance, pp. 303 e ss., Nomos Verlag, Baden-Baden, 2003.

SHLOMAN, Tomer (2025) - Blurring the Lines: How Nation-States and Organized Cybercriminals Are Becoming Alike, *in https://www.trellix.com/blogs/research/blurring-the-lines-how-nation-states-and-cybercriminals-are-becoming-alike/* (acedido em 17/06/2026).

TROPINA, Tatiana (2013) - "*Organized Crime in Cyberspace*", *in* Transnational Organized Crime, pp. 47 e ss., transcript Verlag, Bielefeld, 2013, *in https://www.transcript-publishing.com/media/pdf/93/df/88/oa9783839424957nLX8S8fQiCORI.pdf* (acedido em 17/06/2026).

VENÂNCIO, Pedro Dias (2022) - Lições de Direito do Cibercrime. E da tutela penal de dados pessoais, Editora d'Ideias, Coimbra, 2022.



CYBERLAW

BY CIJIC

PARA ALÉM DO AVATAR - DIGNIDADE HUMANA, IDENTIDADE E COMUNIDADE DIGITAL NA VIDA EXTRAORDINÁRIA DE IBELIN

NUNO POIARES[*]

&

ANA RITA PULGA[†]

* Pós-doutor em Direitos Humanos (IGC/CDH-FDUC), doutor em Sociologia do Direito e Segurança (ISCTE-IUL) e especialista em Direito Penal (Decreto-Lei n.º 206/2009, de 31 de agosto). Mestre em Direito Penal (FDUL) e em Sociologia (U. Évora). Oficial Superior da PSP, professor convidado do ISCPSI, ISCTE-IUL e IGC/CDH-FDUC. Ex-Assessor do Primeiro-Ministro e da Ministra da Administração Interna. Membro do ICPOL-ISCPSI e CIDPCC-FDUL. E-mail: nunopoiares@hotmail.com

† Mestranda e pós-graduada em Ciências Policiais (especialização em Criminologia e Investigação Criminal) pelo Instituto Superior de Ciências Policiais e Segurança Interna, Lisboa. Licenciada em Biologia Humana pela Universidade de Évora. E-mail: rita.pulguinha-94@hotmail.com

RESUMO:

No presente ensaio apresentamos uma reflexão sobre o documentário A Extraordinária Vida de Ibelin (Netflix, 2024), analisando a história de Mats Steen e da sua vida digital como Ibelin no universo de World of Warcraft, assente numa análise sociobiológica e jurídico-normativa. Através deste estudo de caso, foi possível compreender a construção de uma identidade e de uma comunidade no espaço digital, onde uma segunda existência social ultrapassou a dinâmica e as inter-relações humanas na vida física; e como as comunidades virtuais podem funcionar como espaços de inclusão social e reconhecimento, particularmente para pessoas com deficiência ou exclusão estrutural, palco privilegiado para a identidade e a sociabilidade digital, exigindo-se uma reflexão sobre a inclusão e os direitos das pessoas com deficiência e o estatuto jurídico e social das comunidades virtuais.

Palavras-chave: Identidade digital, dignidade humana, comunidades online, mundos virtuais, reconhecimento, avatar.

ABSTRACT:

In this essay, we present an analysis of the documentary The Extraordinary Life of Ibelin (Netflix, 2024), examining the story of Mats Steen and his digital life as Ibelin within the World of Warcraft universe, based on a sociobiological and legal-normative reflection. Through this case study, it has been possible to understand the construction of an identity and a community in the digital space, where a second social existence has surpassed the dynamics and human interrelationships of physical life; and how virtual communities can function as spaces for social inclusion and recognition, particularly for people with disabilities or facing structural exclusion, serving as a privileged arena for digital identity and sociability, thereby requiring reflection on the rights of people with disabilities and inclusion, as well as the legal and social status of virtual communities.

Keywords: Digital Identity, Human Dignity, Online Communities, Virtual Worlds, Recognition, Avatar.

INTRODUÇÃO

O presente ensaio emerge numa sociedade designada como algorítmica[1], líquida[2], em rede[3], figital[4], de risco[5], BANI[6], VUCA[7] e FLUX[8], a par de megatendências, como uma política de destruição generalizada[9], o agravamento das alterações climáticas, a pressão sobre os recursos naturais, a diversificação dos modelos económicos, as evoluções demográficas, um mundo mais urbano e multipolar, a aceleração do desenvolvimento tecnológico, as novas exigências à democracia[10] e a transformação profunda do espaço público tradicional[11], onde se rejeita a ideia de que a sociedade possui um destino histórico inevitável ou que uma autoridade possa deter uma verdade política definitiva[12], aspetos multidimensionais que integram o processo civilizacional (Elias, 2006).

Neste quadro, a experiência humana passa a desenvolver-se num *continuum* em que as fronteiras, entre o presencial e o virtual, se tornam progressivamente mais porosas. É precisamente nesta zona de interseção entre corpo, tecnologia, identidade e sociabilidade, que se situa o documentário *A Extraordinária Vida de Ibelin* (*The Remarkable Life of Ibelin*), realizado por Benjamin Ree e disponibilizado pela Netflix em 2024. A obra recupera a história de Mats Steen (1989-2014), um jovem norueguês diagnosticado com Distrofia Muscular de Duchenne (DMD), doença neuromuscular degenerativa que condicionou a sua mobilidade e o confinou a uma cadeira de rodas. A sua existência, contudo, não pode ser compreendida exclusivamente a partir das limitações impostas pelo corpo. Durante vários anos, Mats construiu uma intensa vida social no universo digital *World of Warcraft*, através

1 Capeller, Pedroso e Santos, 2026.
2 Bauman, 2021.
3 Castells, 2002, 2007.
4 Figital (*phygital*), que une as palavras *physical* (físico) e *digital*, criado em 2007 por Chris Weil, CEO da agência de marketing *Momentum Worlwide*. *In* What is Phygital Marketing? Definition, Examples, Trends | Amazon Ads (acesso em 16 de maio de 2026).
5 Beck, 1998, 2015.
6 Cascio, 2014 (*Brittle, Anxious, Nonlinear, Incomprehensible*).
7 *Volatile, Uncertain, Complex, Ambiguous*.
8 Conceito que descreve o atual ambiente de negócios e trabalho, caracterizado pela mudança constante, volatilidade e rápida evolução tecnológica (*Fast, Liquid, Unknown, and Experimental*).
9 *Munich Security Report* (2026) *in* https://securityconference.org/ (acesso em 20 de junho de 2026).
10 EMP-RePLAN, 2024; Poiares, 2025.
11 Innerarity, 2010.
12 Popper, 1945.

do seu avatar Ibelin Redmoore, integrando a comunidade Starlight e mantendo, paralelamente, o blogue *Musings of life*[13].

A dimensão dessa vida digital apenas se tornou visível após a sua morte, aos 25 anos de idade. Perante as mensagens e manifestações de inúmeros jogadores que procuraram contactar a família, os pais de Mats descobriram uma realidade que desconheciam: aquele jovem que, no espaço físico, vivia limitado pela doença, tinha construído, no espaço virtual, uma extensa rede de relações, amizades e afetos. Durante milhares de horas de interação, Ibelin foi companheiro, amigo, conselheiro e confidente, desempenhando papéis socialmente significativos na vida de outras pessoas. Os arquivos das conversas e os testemunhos dos membros da comunidade permitiram, assim, reconstruir uma dimensão da existência de Mats que permanecera invisível para aqueles que lhe eram mais próximos no mundo físico.

O *caso* de Mats Steen coloca uma questão que ultrapassa a mera análise de uma experiência individual ou de um fenómeno associado aos videojogos. A sua história convida a repensar a distinção entre *first life* e *second life*, sobretudo quando a denominada "segunda vida" digital deixa de constituir um espaço complementar e passa a assumir funções centrais na construção da identidade, das relações sociais e do sentimento de pertença. A vida digital de Ibelin evidencia que uma comunidade virtual não é necessariamente uma comunidade "*menos real*": nela podem estabelecer-se vínculos afetivos, relações de confiança, formas de cooperação, conflitos, reconhecimento e processos de inclusão social com consequências na vida dos indivíduos.

Esta questão adquire particular relevância quando analisada à luz da experiência das pessoas com deficiência. A trajetória de Mats permite problematizar a relação entre deficiência, acessibilidade e participação social, deslocando o olhar de uma conceção centrada apenas nas limitações funcionais do indivíduo para uma perspetiva que considera igualmente as barreiras sociais e ambientais à participação. O espaço digital pode, neste sentido, constituir um território de possibilidades: um ambiente no qual determinadas limitações corporais perdem a centralidade e no qual o indivíduo pode apresentar-se aos outros a partir de capacidades e escolhas que não estão condicionadas pela sua condição física. Ibelin não elimina a deficiência de Mats, mas cria condições para que esta deixe de determinar, de forma absoluta, a possibilidade de estabelecer relações e ocupar um lugar social reconhecido.

13 *Vide* Musings of life (acesso em 16 de maio de 2026).

É neste enquadramento que o presente artigo propõe uma interpretação sociobiológica e jurídica da experiência retratada em *A Extraordinária Vida de Ibelin*, articulando três eixos fundamentais: a) a construção da identidade e da sociabilidade digital; b) os direitos das pessoas com deficiência e os processos de inclusão e reconhecimento; c) e o estatuto jurídico e social das comunidades virtuais. Pretende-se compreender de que modo os ambientes digitais podem produzir formas de pertença, reconhecimento e questionar os limites das categorias jurídicas e sociais tradicionalmente construídas em torno da pessoa, da comunidade e da participação social.

A questão central de investigação pode, assim, ser formulada nos seguintes termos: como interpretar sociobiologicamente a construção de uma identidade e de uma comunidade significativas no espaço digital? Parte-se da hipótese interpretativa de que as comunidades virtuais podem funcionar como espaços efetivos de inclusão social e reconhecimento, particularmente para pessoas sujeitas a situações de deficiência ou de exclusão estrutural. A análise procura, deste modo, ultrapassar uma leitura meramente tecnológica do fenómeno, considerando o digital como espaço de produção de relações sociais, identidades e formas de participação.

Metodologicamente, o artigo assenta numa abordagem interpretativa de natureza sociobiológica, complementada por uma reflexão jurídico-normativa orientada pelos direitos humanos, direitos digitais e pelos princípios da inclusão e da não discriminação. O documentário constitui o objeto empírico central a partir do qual se problematizam as relações entre corpo e identidade, presença e ausência, realidade física e realidade digital, bem como entre autonomia individual, reconhecimento social e pertença comunitária.

A história de Mats Steen/Ibelin permite, *in fine*, colocar uma questão decisiva para a compreensão das sociedades contemporâneas: o que significa estar socialmente presente quando uma parte substancial da vida humana se desenvolve em ambientes digitais? Mais do que uma narrativa sobre um jovem com deficiência que encontrou nos videojogos um espaço de liberdade, *A Extraordinária Vida de Ibelin* constitui um convite à reflexão sobre a natureza da comunidade e da sociabilidade numa sociedade cada vez mais digitalizada. O percurso de Ibelin demonstra que, quando a *second life* se converte numa dimensão substantiva da *first life*, as relações estabelecidas no espaço virtual deixam de ser tratadas como extensões artificiais da realidade e passam a integrar a experiência social e humana dos indivíduos.

I. IDENTIDADE, CORPO E RECONHECIMENTO NA ERA DIGITAL

A relação entre corpo, identidade e reconhecimento constitui uma das questões fundamentais para a compreensão da experiência humana. Durante parte da modernidade, o corpo foi entendido como elemento central da presença social do indivíduo, condicionando a forma como este se apresenta, é percecionado e estabelece relações com os outros. Na sociedade digital, contudo, esta relação torna-se mais complexa. A expansão dos ambientes virtuais, das redes sociais, dos videojogos e das tecnologias imersivas permite que a apresentação do *eu* deixe de estar diretamente dependente da presença física e das características corporais imediatamente visíveis. O espaço digital introduz, assim, novas possibilidades de construção, experimentação e representação da identidade, sem que isso signifique uma rutura com a identidade física ou social do indivíduo[14].

Esta questão adquire particular relevância quando considerada a partir da experiência da deficiência. A Distrofia Muscular de Duchenne (DMD), que marcou a vida de Mats Steen, constitui uma condição neuromuscular progressiva que conduz à perda gradual de força muscular e de mobilidade. Porém, compreender a experiência de Mats apenas a partir das limitações impostas pela doença, significaria adotar uma perspetiva estritamente biomédica da deficiência, insuficiente para explicar a complexidade da sua existência social. A sociologia da deficiência deslocou o foco da deficiência, enquanto atributo exclusivamente individual, para as barreiras sociais, culturais e ambientais que condicionam a participação das pessoas na sociedade.

Neste sentido, importa distinguir entre a condição corporal e a experiência social da deficiência. O corpo de Mats impunha limitações concretas à sua mobilidade e às possibilidades de interação presencial; todavia, muitas das restrições à sua participação social resultavam também da forma como os espaços, as relações e as expetativas sociais são estruturados em torno de determinados padrões de capacidade física. A deficiência não pode, portanto, ser reduzida à incapacidade funcional do indivíduo: ela é igualmente produzida na

14 Embora tenham sido levantadas preocupações sobre o potencial uso indevido de avatares em mundos virtuais para enganar ou influenciar pessoas e grupos vulneráveis, eles também podem ser usados para experiências positivas. Por exemplo, o projeto *Virtual Environments for Rehabilitation of Gender Violence Offender*s (VRespect.Me), financiado pela UE, centra-se na reabilitação dos agressores de violência doméstica, criando um ambiente de empatia, ajudando-os a reconhecer o comportamento violento dos agressores *in* Comissão Europeia: *Shaping Europe´s digital future* (acesso em 15 de junho de 2026).

interação entre essa condição corporal e um ambiente social que pode criar obstáculos à autonomia, à participação e ao reconhecimento.

É neste ponto que a experiência digital de Mats assume um particular significado. O universo de *World of Warcraft* proporcionou-lhe um espaço no qual a sua condição corporal deixou de constituir o principal elemento através do qual os outros o identificavam. Sob a identidade de Ibelin Redmoore, Mats podia deslocar-se, interagir, explorar, estabelecer relações e assumir diferentes papéis sociais sem que a cadeira de rodas ou as limitações motoras determinassem a forma como era reconhecido pelos outros. O avatar não deve, contudo, ser interpretado como uma máscara destinada a ocultar uma realidade corporal. No caso de Ibelin, trata-se antes de uma reconstrução identitária, através da qual Mats expandia as possibilidades de expressão e de participação do seu próprio *eu*.

A este propósito, a conceção dramatúrgica da identidade desenvolvida por Goffman oferece uma chave interpretativa relevante. Para o autor, "*o eu... não é uma coisa orgânica que possui uma localização específica (...) é um efeito dramático que surge difusamente de uma cena que é apresentada*"[15]. A identidade surge, nesta perspetiva, não como uma entidade fixa e isolada, mas como algo que se constrói na interação social e através das formas de apresentação do indivíduo perante os outros. O ambiente digital amplia este processo: existem novas "cenas", novos públicos e novos recursos de representação, entre os quais se encontram os avatares.

O avatar pode, deste modo, ser compreendido como uma forma de mediação entre o indivíduo e o espaço social digital. Trata-se de uma figura virtual tridimensional que pode assumir género, aparência, raça, capacidades e características diferentes das do utilizador que representa. Pode ser humano ou não humano, apresentar ou não uma deficiência e assumir características físicas que não correspondem ao corpo real do indivíduo. Esta plasticidade torna o avatar particularmente relevante para a análise da identidade, uma vez que permite experimentar diferentes formas de apresentação de si e observar como essas representações influenciam o comportamento e a interação social dos utilizadores.

Sherry Turkle antecipou esta transformação ao afirmar que "*estamos a usar a vida no ecrã para pensar as relações, para pensar a identidade (...). Estamos a usar os computadores*

15 Goffman, 1959, p. 252.

para pensar sobre nós próprios"[16]. O espaço digital não é, assim, apenas um instrumento através do qual os indivíduos comunicam; pode constituir igualmente um espaço reflexivo no qual experimentam diferentes dimensões da identidade, exploram relações e redefinem a própria perceção de si. Anos mais tarde, Turkle acrescentaria que "*procuramos na tecnologia formas de estar em relações e, ao mesmo tempo, de nos protegermos delas*"[17], evidenciando a ambivalência das relações mediadas tecnologicamente: a tecnologia aproxima, mas também permite controlar a exposição e gerir a distância perante os outros.

A experiência de Ibelin pode ser interpretada a partir desta ambivalência. O avatar permitia a Mats participar num universo social complexo, mas também oferecia uma mediação entre o seu corpo e os restantes membros da comunidade. Ibelin não era simplesmente uma personagem independente de Mats: era uma expansão de si próprio. Atrás de cada avatar existia uma pessoa real, com sentimentos, expetativas, vulnerabilidades e necessidades de reconhecimento. A identidade virtual não substituía, portanto, a identidade de Mats; acrescentava-lhe uma dimensão social que dificilmente poderia desenvolver-se com a mesma intensidade no espaço físico.

Esta experiência insere-se num fenómeno mais amplo que pode ser designado por *second life*, entendido não necessariamente como uma existência paralela ou fictícia, mas como uma segunda dimensão da existência social. A sociabilidade online permite estabelecer comunidades que ultrapassam fronteiras geográficas, nacionais e corporais. O espaço digital torna-se, assim, um território de pertença, onde os indivíduos podem encontrar reconhecimento e desenvolver formas de capital social através das relações de confiança, reciprocidade, cooperação e amizade que estabelecem.

Os documentários *Second Skin* (2008), *Life 2.0* e *We Met in Virtual Reality* (2022) oferecem diferentes representações desta transformação. Embora incidam sobre experiências distintas, revelam a crescente dificuldade em estabelecer uma separação rígida entre "vida real" e "vida virtual". Nos mundos digitais desenvolvem-se identidades alternativas, relações amorosas, amizades, carreiras, comunidades e experiências de pertença que podem produzir consequências fora do ambiente tecnológico em que tiveram origem. Em *We Met in Virtual Reality*, por exemplo, a realidade virtual surge como espaço de intimidade, ensino, amizade

16 Turkle, 1995, p. 26.
17 Turkle, 2011, p. 12.

e formação comunitária, permitindo que indivíduos ultrapassem determinadas barreiras físicas ou sociais.

A representação cinematográfica da ideia de avatar encontra uma expressão poderosa na trilogia *Avatar* (2009, 2022 e 2025). Embora pertença a um universo ficcional distinto e não constitua uma representação direta das comunidades virtuais contemporâneas, a saga permite explorar a ideia de uma identidade que atravessa o corpo e se relaciona com outro ambiente, outra comunidade e outra forma de existência. A oposição entre a racionalidade instrumental da exploração de Pandora e a relação de interdependência entre os habitantes e o seu ecossistema, introduz uma reflexão sobre a relação entre tecnologia, natureza, comunidade e identidade. Neste sentido, a "segunda vida" representada pela saga não significa abandonar uma existência anterior, mas transformar a compreensão do *eu*, das responsabilidades coletivas e da relação com o mundo.

No caso de Mats Steen, esta transformação adquire uma dimensão concreta. Ibelin Redmoore era um *investigador privado* que percorria Azeroth e que, quando entrava no jogo, podia correr, explorar e deslocar-se com uma liberdade que o seu corpo físico progressivamente lhe negava. O espaço virtual não eliminava a existência da deficiência, mas permitia que ela deixasse de ocupar o centro da relação social. Durante grande parte do tempo, Ibelin podia simplesmente *ser quem desejava ser*.

A dimensão dessa existência torna-se particularmente evidente através dos arquivos da comunidade. Foram recuperadas cerca de 42.000 páginas de mensagens, diálogos, diários de personagens e discussões em fóruns, através das quais foi possível reconstruir uma parte significativa da vida social de Ibelin. Esses registos revelam uma rede de relações que atravessava diferentes países europeus, incluindo a Noruega, a Dinamarca, a Finlândia, a Inglaterra e os Países Baixos. Revelam também sentimentos que não podem ser classificados como meramente virtuais: amizade, paixão, ciúme, raiva, dor e alegria. Entre essas relações encontra-se a ligação afetiva com a avatar Lisette Rumour, uma jovem residente em Breda, nos Países Baixos.

A existência destes sentimentos permite questionar a própria distinção entre relações "reais" e "virtuais". Se uma amizade produz confiança, se uma conversa oferece apoio emocional, se uma relação amorosa produz paixão e sofrimento, ou se uma comunidade preserva a memória de alguém depois da sua morte, em que sentido deverá essa relação ser considerada

menos real por ter ocorrido através de um ecrã? A materialidade do domínio tecnológico não determina, *per se*, a realidade social da relação.

É neste ponto que a experiência de Ibelin evidencia a importância do reconhecimento. Mats não era apenas um utilizador que participava num videojogo: era um membro reconhecido de uma comunidade. Os outros jogadores atribuíam significado à sua presença, procuravam o seu conselho, partilhavam experiências com ele e desenvolveram relações que continuaram a produzir efeitos após a sua morte. A homenagem anual organizada pelos seus amigos constitui uma expressão desse reconhecimento: a comunidade digital preservou a memória de Ibelin e revelou à família de Mats uma dimensão da sua identidade que até então permanecera invisível.

Assim, a experiência de Mats Steen permite compreender o avatar como espaço de expansão identitária, sociabilidade e reconhecimento. A identidade digital não deve ser entendida como uma representação artificial de uma pessoa real, mas como uma dimensão da própria experiência social dessa pessoa. No caso de um indivíduo cuja participação social se encontrava condicionada pelo corpo, o avatar tornou-se um instrumento através do qual foi possível reconstruir possibilidades de presença, relação e pertença.

A história de Ibelin permite, finalmente, inverter a pergunta tradicional sobre a relação entre corpo e identidade. Em vez de perguntar de que modo a condição corporal determina a identidade social, importa também questionar de que modo os ambientes tecnológicos podem criar condições para que uma identidade se desenvolva para além das limitações impostas por determinado contexto físico. A vida digital de Mats demonstra que o corpo continua a existir e a condicionar a experiência humana, mas não constitui o único território a partir do qual a identidade pode ser construída e reconhecida.

Neste sentido, Ibelin não é o contrário de Mats; é uma das formas através das quais Mats pôde existir socialmente. A sua história demonstra que, na era digital, a identidade se constrói cada vez mais através da articulação entre corpo, tecnologia, representação e relação com os outros. É nesta interseção que as comunidades virtuais podem adquirir uma função de inclusão: não porque substituam o mundo físico, mas porque podem criar espaços sociais nos quais indivíduos limitados nas suas possibilidades de participação encontram novas formas de presença e reconhecimento.

II. INCLUSÃO SOCIAL, DEFICIÊNCIA E CIDADANIA DIGITAL

A inclusão social das pessoas com deficiência constitui uma exigência dos direitos humanos e um dos desafios centrais das sociedades contemporâneas. A deficiência não se manifesta apenas através das limitações funcionais decorrentes de uma determinada condição física ou clínica, mas também através das barreiras que a sociedade constrói ou mantém nos seus espaços, relações e representações. Como sintetiza Shakespeare, "*a deficiência é a interação entre o impedimento individual e a restrição social, e não uma ou outra isoladamente*"[18]. Esta perspetiva permite compreender a experiência de Mats Steen para além da Distrofia Muscular de Duchenne: a questão não reside apenas naquilo que o seu corpo podia ou não podia fazer, mas também nas oportunidades que a sociedade lhe oferecia para participar, comunicar, estabelecer relações e ser reconhecido.

O modelo social da deficiência, associado à obra de Oliver, representa uma mudança fundamental nesta compreensão. Para o autor, "*a deficiência é um estado social, não biológico (...). É a sociedade que incapacita as pessoas com impedimentos*"[19]. Esta formulação não pretende negar a dimensão corporal da deficiência, mas evidencia que a incapacidade resulta igualmente da interação entre o indivíduo e um ambiente que pode ser inadequado às suas necessidades: barreiras físicas, comunicacionais, sociais e culturais podem limitar a autonomia e a participação, tanto quanto as próprias características da condição física. No caso de Mats, a progressão da doença restringia a sua mobilidade, mas era sobretudo no domínio da participação social que a experiência digital de Ibelin viria a revelar-se uma possibilidade alternativa de inclusão.

Esta conceção encontra correspondência no modelo dos direitos humanos da deficiência. Quinn e Degener defendem que esse modelo "*foca-se na dignidade inerente do ser humano e, posteriormente, mas apenas se necessário, nas características médicas da pessoa*"[20]. A pessoa não deve, assim, ser definida pela sua deficiência, mas pela sua dignidade, autonomia, capacidade de participação e titularidade de direitos. A questão central passa de "*o que é que esta pessoa consegue fazer?*" para "*que condições devem ser criadas para que esta pessoa possa participar em igualdade de circunstâncias?*".

18 Shakespeare, 2006, p. 2.
19 Oliver, 1990, p. 11.
20 Quinn & Degener, 2002, p. 1.

Esta mudança de paradigma encontra fundamento no sistema internacional e constitucional de proteção dos direitos humanos. A Declaração Universal dos Direitos Humanos (1948) estabelece a igualdade em dignidade e direitos e reconhece a todas as pessoas o direito de participar na vida social, sem discriminação. A Convenção sobre os Direitos das Pessoas com Deficiência, adotada pelas Nações Unidas em 2006 e ratificada por Portugal, aprofunda esta perspetiva, colocando no centro princípios como a dignidade, a autonomia individual, a não discriminação, a participação e inclusão plena na sociedade, a igualdade de oportunidades e a acessibilidade. No ordenamento jurídico português, a Constituição da República Portuguesa, particularmente através dos princípios da dignidade da pessoa humana, da igualdade e da não discriminação e da proteção das pessoas com deficiência, estabelece igualmente uma base jurídico-constitucional para a promoção da participação e da integração social.

Neste contexto, a inclusão digital assume uma importância crescente. Se a participação social constitui uma dimensão dos direitos humanos, e se uma parte cada vez maior da vida social se desenvolve através de tecnologias digitais, o acesso efetivo a esses ambientes não pode ser entendido como uma questão meramente tecnológica. A cidadania contemporânea pressupõe a capacidade de comunicar, participar, criar relações, aceder à informação e desenvolver atividades em ambientes digitais. A exclusão digital pode, por isso, reproduzir ou aprofundar formas de exclusão social já existentes.

É neste ponto que a reflexão de Manuel Castells se revela pertinente. Para o autor, "*a tecnologia não determina a sociedade. Nem a sociedade dita o curso da mudança tecnológica (...) na verdade, o dilema do determinismo tecnológico é provavelmente um falso problema, visto que a tecnologia é a sociedade, e a sociedade não pode ser compreendida ou representada sem as suas ferramentas tecnológicas*"[21]. A tecnologia deve, portanto, ser analisada como parte integrante das relações sociais e não como uma realidade exterior à sociedade. O espaço digital é simultaneamente produto das estruturas sociais existentes e um instrumento capaz de as transformar.

Luciano Floridi aprofunda esta ideia ao afirmar que "*não somos espetadores imóveis e isolados de um espetáculo digital (...) somos organismos informacionais, profundamente*

21 Castells, 1996, p. 5.

integrados num ambiente informacional global"[22]. A vida contemporânea desenvolve-se num ambiente em que o digital e o físico se encontram interligados. Consequentemente, falar de inclusão social implica considerar as condições de participação nesse ambiente informacional. O acesso à tecnologia, porém, não é suficiente: é necessário garantir competências digitais, acessibilidade, segurança e utilização consciente dos recursos tecnológicos[23].

Os estudos desenvolvidos por Poiares (2019a, 2019b) e Bernardino *et al.* (2021), designadamente sobre a necessidade de desenvolver soluções de cibersegurança específicas para a população sénior, evidencia que a vulnerabilidade digital não se limita às pessoas com deficiência. Diferentes grupos podem encontrar obstáculos específicos no ambiente digital, exigindo respostas diferenciadas de proteção e capacitação. Do mesmo modo, Poiares sublinha a importância de promover uma utilização consciente da Internet, particularmente perante utilizadores ou grupos mais vulneráveis[24]. A cidadania digital deve, por isso, combinar participação e proteção: não basta permitir o acesso ao espaço digital; é necessário assegurar que esse acesso ocorre em condições de segurança, autonomia e dignidade.

A experiência de Mats Steen constitui uma ilustração desta possibilidade. Num primeiro momento, *World of Warcraft* pode parecer apenas uma forma de entretenimento. No entanto, a história de Mats demonstra que as relações estabelecidas num ambiente virtual podem adquirir uma profundidade e um significado comparáveis aos das relações desenvolvidas no mundo físico. Para alguém cuja condição corporal impunha fortes restrições à mobilidade e à interação presencial, Azeroth representou um espaço no qual determinadas barreiras deixavam de assumir a mesma importância. Através de Ibelin, Mats podia deslocar-se, comunicar, participar numa comunidade e construir relações sem que a sua deficiência fosse necessariamente o primeiro elemento a definir a interação.

Esta possibilidade pode ser entendida como uma superação simbólica das limitações físicas. O avatar não eliminava as limitações concretas do corpo de Mats; permitia, contudo, criar um contexto social no qual essas limitações não determinavam integralmente a sua participação. Ibelin tornou-se, neste sentido, uma extensão da experiência de Mats e uma forma de exercer uma autonomia que o corpo físico restringia.

22 Floridi, 2013, p. 6.
23 Poiares, 2019a, 2019b; Bernardino et al, 2021.
24 Poiares, 2020, 2024.

A questão da inclusão, porém, não se esgota na experiência individual de Mats. Importa igualmente perguntar quem eram os restantes jogadores e o que procuravam naquele espaço. Se Ibelin encontrou em Azeroth amizade, reconhecimento e pertença, outros utilizadores poderiam encontrar respostas para necessidades sociais, emocionais ou relacionais que não estavam plenamente satisfeitas no mundo físico. O universo digital pode constituir, assim, um espaço de experimentação identitária e de construção comunitária, no qual os indivíduos selecionam, transformam ou ocultam determinadas características da sua identidade.

Surge, assim, uma questão fundamental: até que ponto podemos considerar estas relações uma forma de inclusão social? A inclusão não deve ser reduzida à presença física de um indivíduo num determinado espaço. Participar significa também poder comunicar, estabelecer vínculos, contribuir para uma comunidade, ser ouvido e sentir-se reconhecido. Sob esta perspetiva, as comunidades digitais podem funcionar como espaços de participação social relevantes para pessoas que enfrentam barreiras físicas, geográficas ou comunicacionais.

Todavia, importa evitar uma visão excessivamente otimista ou determinista da tecnologia. O digital pode aproximar pessoas que se encontram geograficamente distantes, mas não garante, por si só, a integração social. É possível estar envolvido em inúmeras conversas e comunidades digitais e, simultaneamente, experimentar a solidão. A quantidade de interações não determina a qualidade dos vínculos. Da mesma forma, a comunicação mediada por tecnologia pode facilitar a expressão emocional de algumas pessoas, mas também pode criar novas formas de vulnerabilidade ou isolamento.

Daqui decorre uma dimensão de responsabilidade afetiva nas relações digitais. Por detrás de cada avatar, perfil ou identidade digital existe uma pessoa real, dotada de sentimentos, expetativas e vulnerabilidades. A natureza virtual da interação não diminui a responsabilidade pelos efeitos produzidos sobre os outros. A ausência, a rejeição ou determinadas formas de comunicação podem ter consequências emocionais reais. O facto de a relação ocorrer através de um ecrã não transforma o outro num objeto virtual.

A história de Mats permite compreender a inclusão digital como uma dimensão contemporânea da cidadania. O seu percurso não deve ser interpretado apenas como a história de um jovem condicionado por uma doença degenerativa, mas também como a história de uma sociedade confrontada com a necessidade de criar novas formas de participação. Ibelin

proporcionou a Mats uma possibilidade de presença social que o seu contexto físico tornava mais difícil, demonstrando que a tecnologia pode constituir uma ferramenta de inclusão quando concebida e utilizada em função da dignidade e da autonomia das pessoas.

A verdadeira questão não é, portanto, saber se a vida digital é "real" ou "virtual", mas compreender que efeitos sociais reais são produzidos através dela. No caso de Mats Steen, esses efeitos foram a amizade, o amor, reconhecimento, pertença e memória. A sua experiência demonstra que a cidadania digital não pode ser separada da cidadania social: ambas pressupõem a possibilidade de estabelecer relações e ser reconhecido como pessoa. O desafio contemporâneo consiste, assim, em assegurar que os ambientes digitais não reproduzem as barreiras existentes no mundo físico, mas contribuem para as ultrapassar, criando espaços de participação efetiva, acessível, segura e inclusiva.

III. COMUNIDADES VIRTUAIS, RECONHECIMENTO JURÍDICO E NOVAS FORMAS DE SOCIABILIDADE

As transformações tecnológicas, das últimas décadas, modificaram as formas através das quais os indivíduos comunicam, constroem relações e participam em comunidades. O desenvolvimento de ambientes digitais imersivos tornou difícil sustentar uma separação absoluta entre relações "reais" e relações "virtuais". Como refere Boellstorff, "*os mundos virtuais são lugares de imaginação que se tornaram realidades sociais, onde as pessoas vivem vidas que são virtualmente humanas através dos seus avatares*"[25]. Esta formulação permite compreender que o caráter virtual do ambiente não elimina a realidade social das experiências que nele ocorrem. Os espaços digitais podem ser ficcionais, mas as pessoas que neles interagem, os sentimentos que experimentam e as consequências das suas relações pertencem à experiência humana.

A história de Mats Steen e da comunidade *Starlight* constitui um exemplo desta realidade. Durante anos, Mats participou no universo de *World of Warcraft* através de Ibelin, construindo uma rede de relações que a sua própria família desconhecia. No seu funeral, o pai recordou uma das suas maiores mágoas: acreditar que o filho, devido às limitações que condicionavam a sua vida física, não teria a oportunidade de experienciar o amor, uma paixão

25 Boellstorff, 2008, p. 4.

ou amizades. Nos dias que se seguiram à morte de Mats, porém, compreendeu que estava enganado. As mensagens e os testemunhos recebidos revelaram uma dimensão da vida do filho: Mats tinha sido relevante em muitas vidas humanas.

A comunidade *Starlight* demonstra, assim, que uma comunidade digital pode constituir um espaço social significativo. Os seus membros não se limitavam a partilhar um ambiente tecnológico ou uma atividade de entretenimento; estabeleciam relações de amizade, confiança, cooperação, conflito, intimidade e pertença. A morte de Ibelin produziu, por isso, um efeito comunitário: os jogadores sentiram a perda de alguém que integrava a sua vida social. A manutenção de homenagens e encontros em memória de Mats demonstra que a comunidade não desapareceu com o encerramento da sessão de jogo. Pelo contrário, a relação entre os seus membros ultrapassou o próprio espaço virtual.

Esta realidade pode ser analisada à luz da teoria do reconhecimento de Axel Honneth. Segundo o autor, "*os sujeitos só se experienciam como membros de pleno direito de uma sociedade se forem mutuamente confirmados em certas formas de reconhecimento social*"[26]. O reconhecimento constitui, nesta perspetiva, uma condição fundamental da integração social e da construção de uma relação positiva do indivíduo consigo próprio. Ser reconhecido significa não apenas ser visto, mas ser considerado relevante por outros, participar em relações de reciprocidade e ocupar um lugar numa comunidade.

A experiência de Ibelin permite compreender concretamente esta ideia. Mats não era um participante anónimo num universo digital composto por milhões de utilizadores. Era confidente, conselheiro e membro de uma comunidade. A importância da sua presença tornou-se evidente quando faleceu. O reconhecimento que recebeu em vida prolongou-se, depois da morte, através da memória construída pelos membros da comunidade. Neste sentido, a experiência de Mats revela que as comunidades virtuais podem funcionar como espaços de reconhecimento social efetivo, sobretudo para pessoas que, por razões físicas, geográficas, sociais ou culturais, encontram dificuldades acrescidas na participação em determinados contextos presenciais.

Esta constatação não significa, contudo, que o espaço digital seja um território socialmente neutro ou juridicamente vazio. As interações digitais produzem consequências concretas e colocam problemas relacionados com a identidade, a privacidade, a dignidade, a

26 Honneth, 1992, p. 11.

responsabilidade e a proteção da personalidade. A liberdade de construir um avatar e de assumir diferentes papéis pode constituir uma importante forma de expressão e experimentação identitária, mas não elimina as consequências dos comportamentos adotados. Surgem, por isso, questões fundamentais: quais são os limites da liberdade de representação no ambiente digital? Em que medida a identidade virtual deve ser protegida juridicamente? E como deve o Direito responder quando comportamentos praticados através de uma identidade digital produzem danos reais?

A reflexão de Lawrence Lessig permanece relevante neste domínio, ao afirmar que "*o código é a lei. O software e o hardware que fazem do ciberespaço o que ele é regulam o ciberespaço tão firmemente quanto qualquer lei escrita por uma legislatura regula um espaço físico*"[27]. O autor evidencia que a regulação do espaço digital não resulta exclusivamente das normas jurídicas tradicionais. A arquitetura tecnológica condiciona aquilo que os utilizadores podem fazer, ver, revelar, ocultar ou controlar. As regras de uma plataforma, os mecanismos de privacidade, as possibilidades de criação de avatares e os sistemas de moderação constituem formas concretas de ordenação da vida digital.

Daqui resulta uma dupla dimensão regulatória. Por um lado, existem normas jurídicas aplicáveis à proteção da dignidade humana, da personalidade, da privacidade, da identidade e da segurança dos indivíduos. Por outro, existe uma arquitetura tecnológica que estabelece as condições práticas de exercício desses direitos. O desafio contemporâneo consiste, assim, em garantir que o desenvolvimento tecnológico não cria espaços onde a proteção da pessoa fique subordinada exclusivamente às regras definidas por plataformas privadas ou pelos códigos informáticos que estruturam a interação.

Neste enquadramento, a Declaração Universal dos Direitos Humanos, a Convenção sobre os Direitos das Pessoas com Deficiência e a Constituição da República Portuguesa oferecem referências essenciais. A dignidade da pessoa humana, a igualdade, a não discriminação, a autonomia e o direito à participação social não deixam de ser relevantes pelo simples facto de uma interação ocorrer através de um ecrã ou de um avatar. Pelo contrário, numa sociedade em que uma parte crescente da existência decorre em ambientes digitais, estes princípios devem ser interpretados de forma capaz de acompanhar as novas modalidades de participação e sociabilidade.

27 Lessig, 1999, p. 6.

A proteção jurídica da participação das pessoas com deficiência encontra fundamento num conjunto articulado de instrumentos internacionais e nacionais que permitem compreender a inclusão digital como uma dimensão contemporânea da efetivação dos direitos fundamentais. A Convenção sobre os Direitos das Pessoas com Deficiência (CDPD), adotada pelas Nações Unidas em 2006 e ratificada por Portugal, estabelece, no artigo 3.º, como princípios gerais, o respeito pela dignidade inerente, a autonomia individual, a não discriminação, a participação e inclusão plena e efetiva na sociedade, a igualdade de oportunidades e a acessibilidade. Esta última dimensão assume particular relevância no contexto das comunidades virtuais, uma vez que o artigo 9.º impõe aos Estados a adoção de medidas destinadas a assegurar, em condições de igualdade com as demais pessoas, o acesso à informação e às comunicações, incluindo as tecnologias e sistemas de informação e comunicação, determinando ainda a promoção do acesso das pessoas com deficiência às novas tecnologias e à Internet e o desenvolvimento de tecnologias e sistemas acessíveis desde a fase inicial da sua conceção. O artigo 19.º reforça esta perspetiva ao reconhecer o igual direito das pessoas com deficiência a viverem de forma independente e a serem incluídas na comunidade, enquanto o artigo 21.º protege a liberdade de expressão e de opinião e o acesso à informação através das formas de comunicação escolhidas pela própria pessoa. Também o artigo 29.º, ao assegurar a participação na vida política e pública, evidencia que a cidadania das pessoas com deficiência deve ser concebida em termos de participação efetiva e não de mera presença formal. No ordenamento jurídico português, estes princípios encontram correspondência na Constituição da República Portuguesa, designadamente no artigo 13.º, que consagra o princípio da igualdade e da igual dignidade social, no artigo 26.º, que reconhece os direitos à identidade pessoal, ao desenvolvimento da personalidade, ao bom nome e reputação, à imagem, à palavra, à reserva da intimidade da vida privada e familiar e à proteção legal contra quaisquer formas de discriminação, e no artigo 71.º, que estabelece que as pessoas com deficiência gozam plenamente dos direitos consignados na Constituição, incumbindo ao Estado promover a sua integração e assegurar a efetiva realização dos seus direitos. Esta proteção é complementada pela Lei n.º 46/2006, de 28 de agosto, que proíbe e pune a discriminação em razão da deficiência e que, no seu artigo 4.º, qualifica como práticas discriminatórias as ações ou omissões, dolosas ou negligentes, que, em razão da deficiência, violem o princípio da igualdade, incluindo expressamente, na alínea m), a adoção de medidas

que limitem o acesso às novas tecnologias. Esta previsão assume especial significado no contexto das comunidades virtuais, pois permite estabelecer uma conexão jurídica direta entre a acessibilidade tecnológica e a igualdade material: a exclusão digital de uma pessoa com deficiência não deve ser analisada como uma insuficiência técnica ou uma dificuldade de utilização, podendo, em determinadas circunstâncias, configurar uma prática discriminatória juridicamente relevante.

Paralelamente, o reconhecimento da relevância social das comunidades virtuais não significa a criação de um espaço juridicamente autónomo ou imune às regras destinadas à proteção da pessoa e dos seus bens jurídicos fundamentais. A Constituição da República Portuguesa protege, no artigo 26.º, dimensões relevantes para a sociabilidade digital, estabelecendo garantias contra a obtenção e utilização abusivas de informações relativas às pessoas e famílias. No âmbito jurídico-penal, esta proteção encontra concretização em diversos tipos legais suscetíveis de aplicação a condutas praticadas através de meios digitais. Os artigos 180.º e 181.º do Código Penal protegem, respetivamente, a honra e consideração através da incriminação da difamação e da injúria; o artigo 154.º-A pune a perseguição reiterada ou o assédio, direto ou indireto, quando a conduta seja adequada a provocar medo ou inquietação ou a prejudicar a liberdade de determinação da vítima; e os artigos 192.º e 199.º asseguram proteção penal perante determinadas formas de devassa da vida privada, gravação ou utilização ilícita de palavras, fotografias ou outros conteúdos. Particularmente relevante para a realidade digital é ainda o artigo 193.º do Código Penal, relativo à devassa por meio de tratamento informático ou de divulgação pública generalizada, cuja redação atual contempla a disseminação através da Internet. Deste modo, a circunstância de uma conduta ser praticada através de um avatar, de uma plataforma de videojogo, de uma rede social ou de outro ambiente virtual não elimina a sua relevância jurídica nem afasta, *per se*, a responsabilidade do agente. A identidade virtual constitui uma forma de representação e de participação social, mas não uma personalidade juridicamente desligada do indivíduo que a utiliza. Por detrás de cada avatar existe uma pessoa titular de direitos e sujeita a deveres, e por detrás de cada interação digital podem existir efeitos concretos sobre a honra, a privacidade, a imagem, a liberdade, a dignidade e o desenvolvimento da personalidade. Assim, as comunidades virtuais devem ser compreendidas como espaços de sociabilidade juridicamente relevantes, pois podem promover a autonomia, a inclusão, reconhecimento e participação, como

evidencia a experiência de Mats Steen, mas estão igualmente sujeitas aos princípios da igualdade, da não discriminação, da proteção da personalidade e da responsabilidade jurídica. A natureza virtual do espaço não transforma, por conseguinte, os comportamentos nele praticados em atos desprovidos de consequências, nem permite conceber a identidade digital como um território de irresponsabilidade ou impunidade.

Esta realidade conduz-nos, inevitavelmente, à questão da personalidade digital. A identidade construída num ambiente virtual pode integrar dimensões relevantes da personalidade de um indivíduo: relações, memórias, comunicações, reputação e formas de expressão. A proteção da pessoa no ambiente digital deve, por isso, considerar não apenas os dados técnicos associados à sua presença online, mas também o significado social e afetivo que essa presença pode assumir. Por detrás de um avatar existe uma pessoa; por detrás de uma identidade digital existem interesses, direitos e relações que podem merecer proteção jurídica.

A morte de Mats Steen evidencia ainda outra dimensão: a do luto digital, da memória e do património afetivo online. A morte física não determina necessariamente o desaparecimento da presença digital. Permanecem mensagens, conversas, publicações, personagens, imagens e outros vestígios de uma existência social. No caso de Mats, os arquivos digitais permitiram reconstruir uma parte substancial de uma vida que a família desconhecia. Os milhares de páginas de mensagens e diálogos revelaram relações, emoções e acontecimentos que constituíam um verdadeiro património afetivo e social.

A memória digital coloca, assim, novos desafios jurídicos e éticos. Quem deve poder aceder às informações deixadas por uma pessoa após a sua morte? Qual é o estatuto das mensagens privadas, das contas, dos conteúdos criados e das identidades digitais? Até que ponto os interesses da família, da comunidade e da memória do falecido devem ser conciliados com a proteção da privacidade e da vontade do próprio indivíduo? Estas questões demonstram que o património digital não se reduz ao seu valor económico: pode possuir um valor relacional, identitário e afetivo.

A experiência de Mats permite igualmente ultrapassar a ideia de que jogar constitui uma atividade separada da vida social. Como refere Taylor, “o jogar não é um espaço separado e limitado, cortado do mundo, mas sim uma interseção complexa onde as culturas online e

offline se encontram e se constroem mutuamente"[28]. A comunidade *Starlight* confirma esta perspetiva. As relações desenvolvidas em *World of Warcraft* não ficaram encerradas em Azeroth: influenciaram vidas concretas, produziram memórias e conduziram, após a morte de Mats, ao encontro entre a sua família e os membros da comunidade digital.

Contudo, a crescente mediação tecnológica das relações humanas conduz também a uma questão crítica: até que ponto a tecnologia pode contribuir para uma desumanização das relações e dos afetos? Se algumas pessoas se sentem mais confortáveis a expressar emoções quando o interlocutor não está associado a um rosto físico, estaremos mais próximos ou estaremos apenas a criar novas formas de distância? A tecnologia pode aproximar pessoas separadas por milhares de quilómetros e, simultaneamente, favorecer formas de afastamento relativamente àqueles que se encontram fisicamente próximos. A comunicação digital facilita a expressão, mas também pode reduzir determinados elementos da presença humana; permite controlar a imagem que apresentamos, mas pode dificultar a exposição da nossa vulnerabilidade. Os ambientes virtuais oferecem liberdade para experimentar identidades, mas também colocam problemas de autenticidade, manipulação, consentimento e responsabilidade.

A liberdade de assumir diferentes papéis deve, por isso, ser compatibilizada com a proteção dos direitos dos outros. É legítimo experimentar identidades e explorar diferentes formas de expressão, mas essa liberdade encontra limites quando a identidade virtual é utilizada para manipular, enganar de forma lesiva, violar a privacidade ou causar danos[29]. A questão torna-se mais complexa quando as fronteiras entre a representação virtual e a identidade pessoal deixam de ser percecionadas. Os comportamentos podem ser praticados através de um avatar, mas os seus efeitos podem atingir pessoas reais.

A experiência de Mats Steen demonstra, contudo, que seria errado compreender o mundo digital apenas como um espaço de alienação ou fuga à realidade. Para algumas pessoas, pode representar uma oportunidade de participação, comunicação, autonomia e pertença. Ibelin permitiu a Mats estabelecer relações, viver experiências, construir memórias e deixar uma marca na vida de outras pessoas, mesmo quando as limitações do seu corpo condicionavam profundamente a sua experiência no mundo físico.

28 Taylor, 2006, p. 19.
29 A este propósito *vide* Fernandes, 2025.

Mats deixou, deste modo, uma marca simultaneamente no mundo digital e no mundo físico. A sua história demonstra que uma relação não perde necessariamente significado por ser mediada por um ecrã. O que determina a sua relevância é a existência de pessoas, emoções, vínculos, responsabilidades e partilhas. Talvez a questão decisiva não seja saber se estamos mais próximos ou mais afastados num mundo cada vez mais digitalizado, mas compreender de que modo utilizamos a tecnologia para criar ou impedir condições de participação, reconhecimento e pertença.

A tecnologia pode constituir uma ponte ou uma barreira. Pode aproximar quem está distante, facilitar a participação de quem encontra obstáculos no espaço físico e permitir a formação de comunidades. Mas também pode aprofundar o isolamento, reproduzir desigualdades e criar novas formas de vulnerabilidade. É por isso que a cidadania digital exige liberdade, responsabilidade, proteção jurídica e ética.

A inclusão social exige mais do que a simples aceitação abstrata da diferença. Exige a capacidade de reconhecer a experiência do outro e de criar condições concretas para que cada indivíduo possa viver com dignidade, autonomia e participação efetiva na sociedade. A empatia não corresponde apenas à tolerância perante as escolhas alheias; implica reconhecer necessidades, vulnerabilidades e expetativas, contribuindo para que o outro possa exercer plenamente a sua condição de membro da comunidade.

No final, a história de Ibelin deixa-nos perante uma reflexão que ultrapassa o universo de *World of Warcraft*. Por detrás de cada avatar existe uma pessoa; por detrás de cada interação digital existe uma dimensão humana; e por detrás de cada identidade construída virtualmente pode existir alguém à procura de uma forma de pertencer. As comunidades virtuais podem, por isso, constituir novos espaços de sociabilidade e reconhecimento que merecem ser compreendidos não como substitutos *inferiores* da vida social, mas como uma das suas expressões contemporâneas.

IV. CONCLUSÃO

Vimos que as comunidades digitais podem desempenhar um papel essencial na inclusão social, reconhecimento e dignidade humana, especialmente para pessoas com deficiência. Há, contudo, diversas implicações jurídicas no contexto da necessidade de aprofundar a proteção da identidade digital, e o reconhecimento das comunidades virtuais como espaços relevantes de socialização e cidadania.

A história de Mats Steen, vivida através do avatar Ibelin, permite-nos questionar uma das distinções mais persistentes da contemporaneidade: a separação entre aquilo que designamos por vida real e a vida virtual. A análise desenvolvida ao longo deste ensaio demonstra que essa fronteira é cada vez mais ténue e sociologicamente menos sustentável. As relações estabelecidas através de um ecrã não são menos humanas, nem os sentimentos, os vínculos, as experiências de pertença ou as formas de reconhecimento que delas resultam deixam de possuir significado. Por detrás de cada avatar existe uma pessoa e, por detrás de cada interação digital, pode existir uma necessidade humana de comunicar, participar, ser reconhecida e pertencer.

O caso de Mats Steen evidencia que os ambientes digitais podem assumir uma função de inclusão social, sobretudo quando as barreiras físicas, geográficas, sociais ou comunicacionais condicionam a participação no mundo presencial. Para Mats, Ibelin não constituiu simplesmente uma personagem de um videojogo, mas uma extensão da sua própria experiência, permitindo-lhe exercer formas de autonomia, sociabilidade e identidade que o seu corpo limitava. A comunidade *Starlight* demonstra, por seu turno, que uma comunidade virtual pode adquirir densidade social e afetiva, produzindo amizades, relações amorosas, solidariedades, conflitos, memórias e reconhecimento. A experiência de Mats confirma, assim, que a participação social não depende exclusivamente da presença física, podendo encontrar nos espaços digitais novas possibilidades de expressão e de pertença.

Esta constatação não deve, contudo, conduzir a uma visão romantizada do digital. A conetividade não equivale necessariamente a integração social e a multiplicação das interações não garante, por si só, a qualidade dos vínculos. O mesmo espaço que pode funcionar como ponte pode transformar-se numa barreira; pode aproximar quem está distante, mas também aprofundar o afastamento de quem já se encontra socialmente isolado.

Acresce que a mediação tecnológica coloca novos desafios relacionados com privacidade, consentimento, autenticidade, manipulação, responsabilidade, proteção da personalidade e fronteiras entre a identidade virtual e a identidade pessoal. A liberdade de construir e experimentar diferentes identidades digitais deve, por isso, coexistir com uma cultura de responsabilidade e com mecanismos efetivos de proteção da pessoa.

Um dos principais desafios identificados reside na necessidade de ultrapassar uma conceção do digital enquanto realidade secundária, meramente lúdica ou desligada das relações humanas. As comunidades virtuais são espaços sociais onde as pessoas vivem experiências significativas e onde podem ser construídas formas de capital social, reconhecimento e cidadania. O reconhecimento desta realidade exige uma reflexão jurídica mais aprofundada sobre a identidade digital, os direitos associados à personalidade virtual, a proteção dos dados e das relações estabelecidas em ambientes imersivos, bem como sobre o destino da memória e do património afetivo digital após a morte. A experiência de Mats Steen demonstra que aquilo que permanece num espaço digital pode possuir um valor emocional e social que ultrapassa a dimensão tecnológica que o suporta.

Neste contexto, importa desenvolver uma agenda de ação assente numa ideia fundamental: a tecnologia deve ser concebida para ampliar a dignidade, a autonomia e a participação das pessoas, e não para substituir ou invisibilizar essas dimensões. Será necessário promover ambientes digitais acessíveis e inclusivos, incorporar princípios de desenho universal e acessibilidade desde a conceção das plataformas; reforçar a literacia digital e relacional e desenvolver mecanismos de proteção atentos às pessoas em situação de maior vulnerabilidade ou exclusão. Paralelamente, torna-se importante estimular uma maior cooperação entre investigadores, decisores políticos, profissionais das áreas sociais e da saúde, indústria tecnológica, comunidades de utilizadores e organizações representativas das pessoas com deficiência.

O futuro poderá levar-nos ainda mais longe: a expansão da realidade virtual e dos ambientes imersivos, da inteligência artificial e de formas cada vez mais sofisticadas de representação através de avatares poderá transformar a maneira como trabalhamos, aprendemos, nos relacionamos e participamos na sociedade. Neste cenário, a questão essencial não é se o mundo digital substituirá o mundo físico, mas que sociedade queremos construir quando as fronteiras entre ambos se tornarem indistintas. A inovação tecnológica poderá permitir que

pessoas limitadas por barreiras físicas encontrem novas formas de mobilidade social, comunicação, aprendizagem, criação e participação. E podemos caminhar para ambientes digitais onde a acessibilidade passe a constituir um princípio estruturante da própria experiência humana.

É nesta perspetiva que Ibelin pode ser entendido não apenas como uma personagem de *World of Warcraft*, mas como uma metáfora do futuro da condição humana numa sociedade cada vez mais digital. O seu percurso mostra-nos que a tecnologia não é, em si mesma, inclusiva ou excludente: são as escolhas humanas, sociais, jurídicas e políticas que determinam a forma como ela será utilizada. O desafio será, por isso, construir uma cultura digital capaz de reconhecer que a pessoa permanece no centro da experiência tecnológica, independentemente do corpo, do avatar ou do espaço através do qual se manifesta.

Em última análise, a extraordinária vida de Ibelin deixa-nos uma pergunta que ultrapassa o universo dos videojogos e interpela a própria sociedade: estamos a criar tecnologias que permitem às pessoas estar mais próximas ou estamos apenas a aprender novas formas de distanciamento? A resposta dependerá da nossa capacidade de transformar inovação em inclusão, conectividade em pertença e representação digital em reconhecimento humano. Se conseguirmos fazê-lo, os mundos virtuais poderão deixar de ser entendidos como espaços de evasão e passar a constituir territórios de cidadania, dignidade e participação. Nesse futuro, talvez a maior inovação tecnológica não seja criar avatares cada vez mais realistas, mas construir comunidades digitais onde nenhuma pessoa tenha de deixar de ser plenamente humana para poder integrá-las.

REFERÊNCIAS

BAUMAN, Z. (2001). *Modernidade Líquida*. Rio de Janeiro: Zahar.

BECK, U. (1998/2013). *La sociedad del riesgo: hacia uma nueva modernidade*. Paidós.

BECK, U. (2015). *Sociedade de Risco Mundial*. Edições 70.

BERNARDINO, Iolanda; BAPTISTA, Ricardo José Vieira; ALMEIDA, José Manuel Emiliano Bidarra de; SÃO MAMEDE, José Henrique Pereira (2021), "*Serious Games for seniors: Learning safe behaviors on the web : Position paper*" *16th Iberian Conference on Information Systems and Technologies (CISTI)*, Portugal, pp. 1-4, doi: 10.23919/CISTI52073.2021.9476310.

BOELLSTORFF, T. (2008). *Coming of age in Second Life: An anthropologist explores the virtually human*. Princeton University Press.

CAPELLER, W., PEDROSO, J. e SANTOS, A. (2026). *Uma Teoria Sociojurídica da Sociedade Algorítmica: Perspetivas Plurais Críticas*. Edições Almedina.

CASCIO, J. (2014). *A New Social Operating System for the Age of Resilience*, Institute for Ethics and Emerging Technologies.

CASTELLS, M. (1996). *The rise of the network society*. Blackwell Publishers.

CASTELLS, M. (2002). *A Era da Informação: Economia, Sociedade e Cultura, I - A Sociedade em Rede*. Fundação Calouste Gulbenkian.

CASTELLS, M. (2007). *A Sociedade em Rede - A Era da Informação: Economia, Sociedade e Cultura, I – A Sociedade* (3.ª ed.). Fundação Calouste Gulbenkian.

CONSTITUIÇÃO DA REPÚBLICA PORTUGUESA – aprovada em 2 de abril de 1976.

CONVENÇÃO INTERNACIONAL SOBRE ELIMINAÇÃO DE TODAS AS FORMAS DE DISCRIMINAÇÃO SOCIAL. Disponível em: http://gddc.ministeriopublico.pt/sites/default/files/documentos/instrumentos/prev_discriminacao_convencao_internacional_elim_formas_disc_racial.pdf

CONVENÇÃO SOBRE OS DIREITOS DAS PESSOAS COM DEFICIÊNCIA. Disponível em: https://gddc.ministeriopublico.pt/sites/default/files/documentos/instrumentos/pessoas_deficiencia_convencao_sobre_direitos_pessoas_com_deficiencia.pdf

DECLARAÇÃO UNIVERSAL DOS DIREITOS HUMANOS. Disponível em: https://dre.pt/dre/geral/legislacao-relevante/declaracao-universal-direitos-humanos

DECRETO-LEI n.º 400/82, de 23 de setembro - aprova o Código Penal, Diário da República Eletrónico n.º 221/1982, I, 1982-09-23, pp. 2-64.

ELIAS, N. (2006). *O processo civilizacional. Investigações sociogenéticas e psicogenéticas* (2.ª ed.). Publicações Dom Quixote.

FERNANDES, R. (2025). *Outlets de Desinformação na Nova Geopolítica Digital*. Lisboa: ICPOL-ISCPSI.

FLORIDI, L. (2013). *The ethics of information*. Oxford University Press.

GOFFMAN, E. (1959). *The presentation of self in everyday life*. Anchor Books.

HONNETH, A. (1992). *Kampf um Anerkennung: Zur moralischen Grammatik sozialer Konflikte*. Suhrkamp Verlag.

INNERARITY, D. (2010). *O Novo Espaço Público*. Teorema.

LEI n.º 46/2006, de 28 de agosto. Proíbe e pune a discriminação em razão da deficiência e da existência de risco agravado de saúde.

LESSING, L. (1999). *Code and other laws of cyberspace*. Basic Books.

MUNICH SECURITY REPORT (2026) *in* https://securityconference.org/ (acesso em 20 de março de 2026).

OLIVER, M. (1990). *The politics of disablement: A sociological approach*. Macmillan Education.

POIARES, N. (2019a). "*A cibersegurança à luz da criminologia moderna*", *Cyberlaw by CIJIC,* 7. CIJIC-FDUL.

POIARES, N. (2019b). "*Cibersegurança, literacia e resiliência digital dos idosos*", *Anuário Janus 2018-2019: A dimensão externa da segurança interna*, 19, 118-119. Observare-UAL.

POIARES, N. (2020). "*Violência doméstica e redes sociais: a proteção jurídico-penal da vida privada na internet*", *Cyberlaw by CIJIC*, 10. CIJIC-FDUL.

POIARES, N. (2024). *Polícia e Direitos Humanos: Multiculturalismo, Género, Saúde Mental e LGBTQIA+* (2.ª ed.). Edições Almedina.

POIARES, N. (2025). A Governança da Segurança Interna na Europa no Século XXI: Estado, Processo Civilizacional e Sociedade de Risco, *De Legibus – Revista de Direito*, 9, 169-187. CEAD Francisco Suárez.

POPPER, K. (1945). *The Open Society and its enemies*, I. George Routledge & Sons Ltd. Disponível em: The Open Society and its Enemies: The Spell of Plato, Vol. 1, 1st ed. (antilogicalism.com) (acesso em 16 junho 2026).

QUINN, G., & DEGENER, T. (2002). *Human rights and disability*. UNESCO.

REDE DE SERVIÇOS DE PLANEAMENTO E PROSPETIVA DA ADMINISTRAÇÃO PÚBLICA/EQUIPA MULTISSETORIAL DE PROSPETIVA (2024). *Megatendências 2050. O mundo em mudança: impactos em Portugal - Uma breve introdução*. RePLAN/EMP. https://www.planapp.gov.pt/wp-content/uploads/2024/03/MegatendenciasBreveIntro_RePLAN.pdf

SHAKESPEARE, T. (2006). *Disability rights and wrongs*.

TAYLOR, T. L. (2006). *Play between worlds: Exploring online game culture*. The MIT Press. Cambridge, Massachusetts. London, England.

TURKLE, S. (1995). *Life on the screen: Identity in the age of the internet*. Simon & Schuster.

TURKLE, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.



OS CÓDIGOS DE CONDUTA EM MATÉRIA DE PROTEÇÃO DE DADOS

JOEL A. ALVES *

* Professor Auxiliar Convidado da Escola de Direito da Universidade do Minho. Professor Adjunto Convidado da Escola Superior de Tecnologia e Gestão da Universidade Técnica do Porto. Investigador Integrado do JusGov- Centro de Investigação em Justiça e Governação. Of Counsel da MFP - Sociedade de Advogados.

RESUMO:

O presente artigo encerra dois objetivos fundamentais: (i) a montante, contribuir para uma maior clarificação quanto à disciplina jurídica aplicável aos Códigos de Conduta em matéria de proteção de dados, facilitando a sua compreensão por parte de todos aqueles que possam ter interesse na sua elaboração ou adesão; (ii) a jusante, demonstrar o imenso potencial associado a tais instrumentos, apelando a que estes sejam objeto de um maior aproveitamento por parte de responsáveis pelo tratamento e subcontratantes.

Palavras-chave: Códigos de Conduta; soft law; RGPD; proteção de dados; responsabilidade.

ABSTRACT:

This article has fundamental objectives: (i) at the outset, to contribute to greater clarity regarding the legal framework applicable to Codes of Conduct in the field of data protection, facilitating their understanding by all those who may have an interest in their development or adherence; and (ii) ultimately, to demonstrate the immense potential associated with such instruments, calling for their greater use by controllers and processors.

Keywords: Codes of Conduct; soft law; GDPR; data protection; accountability.

I. ENQUADRAMENTO [1]

Conquanto já gozassem de previsão expressa na Diretiva 95/46/CE[2], os Códigos de Conduta vieram assumir um renovado protagonismo no Regulamento Geral sobre a Proteção de Dados[3]. Reflexo disso é o facto de, na anterior Diretiva, a disciplina relativa a estes instrumentos se esgotar num único preceito - o art. 27.º, composto por três números; cenário que nitidamente contrasta com o do atual Regulamento, o qual, além de lhes dedicar duas extensas e detalhadas disposições específicas - o art. 40.º, sob a epígrafe "Códigos de Conduta", composto por 11 números; e o art. 41.º, intitulado "Supervisão dos Códigos de Conduta aprovados", composto por 6 números - contém, ainda, uma série de referências remissivas para tais instrumentos, porventura tão ou mais importantes.

Ora, é pois, à luz desta aparente revolução jurídica, que se impõe «olhar com novos olhos» para os referidos Códigos, procurando ofertar resposta a interrogações tais como as seguintes: (i) o que são Códigos de Conduta?; (ii) quais as funções que o RGPD atribui a estes instrumentos?; (iii) quais as vantagens que responsáveis pelo tratamento e subcontratantes

1 O presente artigo resulta, com adaptações, atualizações e desenvolvimentos, das considerações originalmente publicadas in ALVES, Joel A. / FONSECA, Isabel Celeste M., "*Smart Cities e Proteção de Dados Pessoais: por um Código de Conduta para uma Transição Digital Local Data Protection Friendly*", in Cidades inteligentes e direito, governação Digital e direitos: mapeamento de cidades e código de conduta (coord. Isabel Celeste M. Fonseca), Almedina, Coimbra, 2023, pp. 317-336. As considerações em causa correspondem ao contributo do autor para o citado escrito.

(2) Cfr. art. 27.º da Diretiva 95/46/CE do Parlamento Europeu e do Conselho, de 24 de Outubro de 1995, relativa à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados, onde se lia o seguinte: "*1. Os Estados-membros e a Comissão promoverão a elaboração de Códigos de Conduta destinados a contribuir, em função das características dos diferentes sectores, para a boa execução das disposições nacionais tomadas pelos Estados-membros nos termos da presente directiva. / 2. Os Estados-membros estabelecerão que as associações profissionais e as outras organizações representativas de outras categorias de responsáveis pelo tratamento que tenham elaborado projectos de códigos nacionais ou que tencionem alterar ou prorrogar códigos nacionais existentes, podem submetê-los à apreciação das autoridades nacionais. Os Estados-membros estabelecerão que essas autoridades se certificarão, nomeadamente, da conformidade dos projectos que lhe são apresentados com as disposições nacionais tomadas nos termos da presente directiva. Se o considerarem oportuno, as autoridades solicitarão a opinião das pessoas em causa ou dos seus representantes. / 3. Os projectos de códigos comunitários, assim como as alterações ou prorrogações de códigos comunitários existentes, poderão ser submetidos ao grupo referido no artigo 29º O grupo pronunicar-se-á, nomeadamente, quanto à conformidade dos projectos submetidos à sua apreciação com as disposições nacionais adoptadas em aplicação da presente directiva. Se o considerar oportuno, solicitará a opinião das pessoas em causa ou dos seus representantes. A Comissão pode garantir uma publicidade adequada dos códigos aprovados pelo grupo*".

(3) Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho de 27 de abril de 2016 relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados, doravante abreviadamente designado pela sigla «RGPD» ou simplesmente por «Regulamento».

poderão retirar da sua adesão?; (iv) quais os requisitos de admissibilidade e aprovação a que estes códigos estão sujeitos e (v) qual o conteúdo que estes poderão incluir?

No presente artigo procurar-se-á responder a estas e outras questões, tendo em vista dois objetivos primaciais: (i) a montante, contribuir para uma maior clarificação quanto à disciplina jurídica aplicável aos Códigos de Conduta, facilitando a sua compreensão por parte de todos aqueles que possam ter interesse na sua elaboração ou adesão; (ii) a jusante, demonstrar o imenso potencial associado a tais instrumentos, apelando a que estes sejam objeto de um maior aproveitamento por parte de responsáveis pelo tratamento e subcontratantes.

II. O QUE SÃO?

À semelhança do que já que acontecia no quadro da Diretiva 95/46/CE, o RGPD não avança, ao longo dos seus 99 artigos e 173 considerandos, quaisquer pistas sobre o que deve entender-se por um «código de conduta». Sem embargo, o Comité Europeu para a Proteção de Dados teve já oportunidade de esclarecer que, para efeitos do Regulamento, podem estes códigos ser definidos enquanto "*instrumentos de responsabilização voluntários* que estabelecem *regras específicas* em matéria de proteção de dados para categorias de responsáveis pelo tratamento de dados e de subcontratantes"[4] (*destaque nosso*).

Da noção apresentada extraem-se, portanto, duas conclusões fundamentais:

a) Em primeiro lugar, os Códigos de Conduta a que o RGPD alude – designadamente, no seu art. 40.º – constituem instrumentos dotados de uma *vinculatividade restringida*[5]. Isto porque, como a própria expressão "*instrumentos de responsabilização voluntários"* indica, a decisão de aderir a um determinado código – e por consequência, de ficar adstrito às disposições nele contidas – encontra-se na

4 Cfr. Comité Europeu para a Proteção de Dados, "Diretrizes 1/2019 relativas aos Códigos de Conduta e aos Organismos de Supervisão ao abrigo do Regulamento (UE) 2016/679", versão 2.0, 4 de junho de 2019, p. 7.

5 Cfr. PÉREZ, María Mercedes Serrano, "Los Códigos de Conducta (Comentario al Artículo 40 RGPD y al Artículo 38 y Disposición Transitoria Segunda LOPDP", in *Comentario al Reglamento General de Protección de Datos Personales y Garantía de los Derechos Digitales* (coord. Antonio Trancoso Reigada), Tomo I, Civitas / Thomson Reuters, Navarra, 2021, p. 2387.

disponibilidade dos seus potenciais destinatários[6], podendo estes livremente optar por não o fazer, sem que daí decorram quaisquer consequências de relevo[7]. Tanto assim é, que a própria elaboração dos códigos anteriormente mencionados se demonstra facultativa[8], limitando-se o Regulamento a estabelecer que os Estados-Membros, as autoridades de controlo, o Comité Europeu para a Proteção de Dados e a Comissão devem tomar providências no sentido de fomentar a sua elaboração, por parte de associações e outros organismos representantes de categorias de responsáveis pelo tratamento ou de subcontratantes[9];

b) Em segundo lugar, os códigos anteriormente referidos assumem uma natureza *complementar* face ao RGPD[10]. O que significa que os mesmos podem concretizar as disposições de tal diploma, criando "*regras [mais] específicas em matéria de*

6 Nesse sentido, cfr., por todos, KAMARA, Irene, Anotação ao artigo 40.º, in *The EU General Data Protection Regulation (GDPR): a Commentary* (org. Cristopher Kuner, Lee A. Bygrave e Cristopher Docksey), Oxford University Press, Oxford, 2020, p. 721.

7 Não obstante, o mesmo já não sucederá a partir da sua decisão de adesão a um tal Código. É que, nos termos do art. 40.º, n.º 4, do RGPD, um código de conduta deve necessariamente prever mecanismos que permitam "a supervisão obrigatória do cumprimento das suas disposições por parte dos responsáveis pelo tratamento ou subcontratantes que se comprometam a aplicá-lo". Supervisão essa, que deve ser realizada "por um organismo que tenha um nível adequado de competência relativamente ao objeto do código e esteja acreditado para o efeito pela autoridade de controlo competente" (art. 41.º, n.º 1); cabendo ao mesmo tomar "sob reserva das garantias adequadas, as medidas que forem adequadas em caso de violações do código por um responsável pelo tratamento ou por um subcontratante, incluindo a suspensão ou exclusão desse responsável ou subcontratante do código (art. 41.º, n.º 4). Ademais, importa levar em conta que a eventual violação das disposições de um código de conduta pode ainda constituir uma *circunstância agravante*, aquando da decisão de uma autoridade de controlo aplicar um coima, e decidir sobre o seu montante, em caso do eventual incumprimento de um determinado responsável pelo tratamento ou subcontratante com certos princípios e regras previstos no RGPD.

8 O art. 40.º, n.º 2, do RGPD é, a este propósito, elucidativo: "as associações e outros organismos representantes de categorias de responsáveis pelo tratamento ou de subcontratantes *podem elaborar Códigos de Conduta*, alterar ou aditar a esses códigos, a fim de especificar a aplicação do presente regulamento" (destaque nosso). Circunstância que parece não deixar dúvidas quanto à inexistência de qualquer obrigação jurídica, relativamente à feitura deste tipo de instrumentos.

9 Assim decorre do art. 40.º, n.º 1, do RGPD; preceito nos termos do qual se estabelece que *"[o]s Estados-Membros, as autoridades de controlo, o Comité e a Comissão promovem a elaboração de Códigos de Conduta* destinados a contribuir para a correta aplicação do presente regulamento, tendo em conta as características dos diferentes setores de tratamento e as necessidades específicas das micro, pequenas e médias empresas" (*sublinhado nosso*). A mesma ideia parece igualmente encontrar respaldo no Considerando 98 do Regulamento, onde se reforça expressamente a ideia de que "*[a]s associações ou outras entidades que representem categorias de responsáveis pelo tratamento ou de subcontratantes deverão ser incentivadas a elaborar Códigos de Conduta*, no respeito do presente regulamento, com vista a facilitar a sua aplicação efetiva, tendo em conta as características específicas do tratamento efetuado em determinados setores e as necessidades específicas das micro, pequenas e médias empresas" (*sublinhado nosso*).

10 Cfr. HERGUETA, Julián Prieto, "Los Códigos de Conducta (Comentario a los Artículos 40 y 41 RGPD y a los Artículos 38 y Disposición Transitoria Segunda LOPDP", *in Comentario al Reglamento General de Protección de Datos Personales y Garantía de los Derechos Digitales* (coord. Antonio Trancoso Reigada), Tomo I, Civitas / Thomson Reuters, Navarra, 2021, p. 2364.

proteção de dados pessoais", mas jamais substituí-las ou contrariá-las[11]. Donde, a eventual adesão a um código de conduta, naturalmente, também não isente responsáveis pelo tratamento e subcontratantes do cumprimento com as obrigações que se lhe colocam ao abrigo do próprio Regulamento[12] – muito embora, como se verá adiante, possa tornar essa tarefa deveras mais simples e eficaz.

III. PARA QUE SERVEM?

Dito isto, importa ter presente que, do ponto de vista jurídico, o RGPD reconhece aos Códigos de Conduta em matéria de proteção de dados pessoais duas principais *funções* ou *finalidades*.

a) Função de facilitação e efetivação do cumprimento

Assim - e no que concerne à primeira função –, o art. 40.º, n.º 1, do RGPD afigura-se por demais elucidativo: os Códigos de Conduta destinam-se "*a contribuir para a correta aplicação do presente regulamento,* tendo em conta as características dos diferentes setores de tratamento e as necessidades específicas das micro, pequenas e médias empresas" (destaque nosso). Complementarmente, esclarece o considerando 98 do mesmo diploma que tais instrumentos visam ainda "*facilitar a sua aplicação efetiva,* tendo em conta as características específicas do tratamento efetuado em determinados setores e as necessidades específicas das micro, pequenas e médias empresa" (*nosso destaque*).

Para o efeito, prevê o legislador a possibilidade dos códigos em questão serem utilizados, por parte de associações e outros organismos representantes de categorias de responsáveis pelo tratamento e subcontratantes, com vista à *especificação* do RGPD em determinados domínios[13].

---

11 Cfr. PÉREZ, María Mercedes Serrano, "Los Códigos de Conducta (Comentario al Artículo 40 RGPD y al Artículo 38 y Disposición Transitoria Segunda LOPDP", cit., p. 2388.

12 Em sentido idêntico, recordando que o cumprimento com um Código de Conduta "não prejudica o cumprimento do RGPD", cfr. GONÇALVES, Carlos Jorge, Anotação ao artigo 40.º do RGPD, in *Comentário ao Regulamento Geral de Proteção de Dados* (coord. Alexandre Sousa Pinheiro), Almedina, Coimbra, 2018, p. 485.

13 Cfr. art. 40.º, n.º 2, do RGPD. O preceito em causa elenca, aliás, um conjunto de matérias suscetíveis de serem concretizados ao nível de um Código de Conduta, quais sejam: (i) o tratamento equitativo e transparente; (ii) os legítimos interesses dos responsáveis pelo tratamento em contextos específicos; (iii) a recolha de dados pessoais; (iv) a pseudonimização de dados pessoais; (v) a informação prestada ao público e aos titulares dos

O que se pretende é, pois, que tais instrumentos possam *concretizar* e *aclarar* os requisitos materiais e organizacionais, de carácter geral e abstrato, previstos no Regulamento[14], fornecendo orientações *práticas* e *precisas* quanto à forma como estas devem ser aplicadas no contexto de um concreto setor e das atividades de tratamento nele desenvolvidas[15]. Objetivo esse que pode ser alcançado através do recurso a uma terminologia única e pertinente para esse mesmo setor[16]. Ou ainda, mediante a apresentação de cenários concretos e exemplos específicos de «melhores práticas»[17] – apenas para citar as vias mais paradigmáticas.

b) Função de demonstração do cumprimento

Acresce que, em complemento do exposto, o RGPD veio conferir aos Códigos de Conduta uma nova e importante função. Referimo-nos, em concreto, à possibilidade de a adesão voluntária a tais instrumentos, por parte de responsáveis pelo tratamento e subcontratantes, poder funcionar como um *elemento evidenciador* do seu cumprimento com várias obrigações previstas no Regulamento.

i) O novo modelo de proteção de dados pessoais europeu

Para que devidamente se compreenda a relevância de tal inovação, importa, porém, começar por recordar a profunda mudança de paradigma operada pelo RGPD no que diz

dados; (vi) o exercício dos direitos dos titulares dos dados; (vi) as informações prestadas às crianças e a sua proteção, e o modo pelo qual o consentimento do titular das responsabilidades parentais da criança deve ser obtido; (vii) as medidas e procedimentos a que aludem os arts. 24.º (Responsabilidade do responsável pelo tratamento) e 25.º (Proteção de dados desde a conceção e por defeito) do Regulamento, bem como as medidas destinas a garantir a segurança do tratamento, referidas no seu art. 30.º; (viii) a notificação de violações de dados pessoais às autoridades de controlo e a comunicação dessas violações de dados pessoais aos titulares dos dados; (ix) a transferência de dados pessoais para países terceiros ou organizações internacionais; e (x) as ações extrajudiciais e outros procedimentos de resolução de litígios entre os responsáveis pelo tratamento e os titulares dos dados em relação ao tratamento. Sem embargo, ponto é que a Doutrina tem alertado para a circunstância de tal listagem não revestir um carácter *fechado* e *exaustivo*, mas antes meramente *exemplificativo.* Daí que, em tese, um Código possa abordar não só os domínios naquela mencionados, com também quaisquer outros regulados ao nível do Regulamento – como, de resto, melhor teremos oportunidade de ver com maior acuidade adiante.

14 Cfr. VOIGT, Paul / VON DEM BUSSCHE, Axel, *The EU General Data Protection Regulation (GDPR): a practical guide*, Springer, Cham, 2017, pp. 71 e 73.

15 Cfr. Comité Europeu para a Proteção de Dados, "Diretrizes 1/2019 relativas aos Códigos de Conduta e aos Organismos de Supervisão ao abrigo do Regulamento (UE) 2016/679", cit, p. 16.

16 Idem, p. 17.

17 Idem, ibidem.

respeito à abordagem regulatória das atividades de tratamento de dados pessoais desenvolvidas ao nível da União Europeia e dos seus Estados-Membros.

Assim, se até 25 de Maio de 2018 - data a partir do qual o Regulamento se tornou diretamente aplicável[18] - a Diretiva 95/46/CE estabelecia uma *obrigação geral de notificação* às autoridades de controlo (em determinados casos, complementada pela sujeição a um procedimento destinado à emissão ou recusa de uma *autorização administrativa*) sempre que uma entidade pretendesse tratar dados pessoais[19], a verdade é que, após esse momento, tudo mudou, passando tal obrigação a verificar-se apenas em situações muito pontuais[20].

Sucede que, enquanto contrapartida pela eliminação do referido encargo, veio o RGPD introduzir um conjunto de novas imposições, destinadas a transferir o ónus de verificação prévia do cumprimento dos pertinentes princípios e regras em matéria de proteção de dados pessoais, anteriormente cometido à entidade *reguladora*, para a esfera das próprias entidades *reguladas*[21]. Imposições essas, de entre as quais avulta o princípio axial, nos termos do qual se faz impender sobre os responsáveis pelo tratamento o dever de implementar todas as medidas técnicas e organizativas que se afigurem adequadas, de acordo com uma abordagem baseada no risco[22], para *garantir* e a todo o tempo poder *demonstrar* a respetiva conformidade com as disposições do Regulamento (*princípio da responsabilidade*)[23].

18 Cfr. art. 99.º, n.º 2, do RGPD.

19) Sobre a obrigação em causa e as suas respetivas projeções ao nível da recentemente revogada Lei n.º 67/98, de 26 de outubro, cfr., por todos, CASTRO, Catarina Sarmento e, *Direito da Informática, Privacidade e Dados Pessoais: a Propósito da Legalização de Tratamentos de Dados Pessoais (Incluindo Videovigilância, Telecomunicações e Internet), por Entidades Públicas e por Entidades Privadas, e da sua Comunicação e Acesso*, Almedina, Coimbra, 2005, pp. 65 e ss..

20) Para mais desenvolvimentos, cfr. o nosso ALVES, Joel A., *O Novo Modelo de Proteção de Dados Pessoais Europeu: Da heterorregulação à autorregulação publicamente regulada*, pp. 107 e ss.

21) Cfr. CALVÃO, Filipa Urbano, "Garantia de direitos: a proteção dos dados pessoais perante os desafios tecnológicos", in *Garantia de Direitos e Regulação: Perspetivas de Direito Administrativo* (coord. Carla Amado Gomes et al.), AAFDL Editora, Lisboa, 2020, p. 234.

22) Sobre o conceito, cfr., entre outros, LOPES, Teresa Vale, "Responsabilidade e Governação das Empresas no Âmbito do Novo Regulamento sobre a Proteção de Dados", in *Anuário da Proteção de Dados 2018* (coord. Francisco Pereira Coutinho e Graça Canto Moniz), Universidade Nova de Lisboa, Lisboa, 2018, pp. 50 e ss.

23) Cfr. art. 5.º, n.º 2, do RGPD, lido em conjugação com o art. 24.º, n.º 1, do mesmo diploma. Para mais desenvolvimentos, veja-se, por todos, o nosso ALVES, Joel A., *Administração Pública e Dados Pessoais: da necessidade do tratamento ao imperativo da proteção*, Almedina, Coimbra, 2026, pp. 211 e ss.

ii) Manifestações específicas

Ora, é, neste contexto que os Códigos de Conduta adquirem um renovado protagonismo, apresentando-se como um instrumento da maior utilidade para todos aqueles que se veem compelidos a ofertar concretização a tão exigente desiderato.

E é assim, pois que, nos termos do art. 24.º, n.º 3, do RGPD, "*o cumprimento de Códigos de Conduta aprovados conforme referido no artigo 40.º* [do Regulamento] (...) *pode ser utilizado como elemento para demonstrar o cumprimento das obrigações do responsável pelo tratamento*" (destaque nosso), funcionando como *indício de prova*[24] da sua conformidade com uma série de exigências materiais e formais, decorrentes do princípio a que *supra* fizemos referência.

Nesta senda, é desde logo paradigmático o disposto no art. 32.º, n.º 3, do RGPD, no qual se estipula que o cumprimento com este tipo de instrumentos pode ser utilizado como elemento para demonstrar que o responsável pelo tratamento dispõe de medidas técnicas e organizativas adequadas para assegurar a segurança dos dados pessoais por si tratados face aos riscos em presença, garantindo a proteção destes últimos contra o seu tratamento não autorizado ou ilícito, bem como contra a sua perda, destruição ou danificação acidental[25].

Por outro lado, merece igualmente nota o art. 35.º, n.º 8, do Regulamento, onde se prescreve que, "ao avaliar o impacto das operações de tratamento efetuadas pelos responsáveis pelo tratamento ou pelos subcontratantes, em especial para efeitos de uma avaliação de impacto sobre a proteção de dados, é tido na devida conta o cumprimento dos Códigos de Conduta

---

24) Nesse sentido, cfr. GÓMEZ, Alberto Díaz-Romeral, "Los códigos de conducta en el Reglamento General de Protección de Datos", in *Reglamento General de Protección de Datos: hacia un nuevo modelo europeo de privacidade* (org. Miguel Álvarez Caro / Miguel Recio Gayo), Editorial Reus, Madrid, 2016, pp. 396 e ss.

25) Recorde-se que, nos termos do art. 5.º, n.º 1, al. f), do RGPD, incumbe ao responsável pelo tratamento garantir que os dados pessoais sob a sua responsabilidade são "tratados de uma forma que garanta a sua segurança, incluindo a proteção contra o seu tratamento não autorizado ou ilícito e contra a sua perda, destruição ou danificação acidental, adotando as medidas técnicas ou organizativas adequadas" (*princípio da segurança*). Imposição que surge melhor concretizada no art. 32.º, n.º 1, do Regulamento, onde se estabelece que "tendo em conta as técnicas mais avançadas, os custos de aplicação e a natureza, o âmbito, o contexto e as finalidades do tratamento, bem como os riscos, de probabilidade e gravidade variável, para os direitos e liberdades das pessoas singulares, *o responsável pelo tratamento e o subcontratante aplicam as medidas técnicas e organizativas adequadas para assegurar um nível de segurança adequado ao risco*, incluindo, consoante o que for adequado: a) a pseudonimização e a cifragem dos dados pessoais; b) a capacidade de assegurar a confidencialidade, integridade, disponibilidade e resiliência permanentes dos sistemas e dos serviços de tratamento; c) a capacidade de restabelecer a disponibilidade e o acesso aos dados pessoais de forma atempada no caso de um incidente físico ou técnico; d) um processo para testar, apreciar e avaliar regularmente a eficácia das medidas técnicas e organizativas para garantir a segurança do tratamento" (*nosso destaque*). Para mais desenvolvimentos, veja-se, por todos, o nosso ALVES, Joel A., *Administração Pública e Dados Pessoais: da necessidade do tratamento ao imperativo da proteção*, cit., pp. 210 e ss.

aprovados a que se refere o artigo 40.º por parte desses responsáveis ou subcontratantes". O que equivale a dizer que o cumprimento com um tal Código pode não só auxiliar os respetivos aderentes a mais facilmente identificarem e avaliarem os riscos associados às atividades de tratamento por si desenvolvidas, como também a colocarem em prática as medidas mitigadoras mais adequadas para lhes fazer frente[26].

Acresce que, de acordo com o art. 28.º, n.º 5, do RGPD, podem estes instrumentos também ser utilizados para demonstrar que um dado subcontratante oferece garantias suficientes – nomeadamente, em termos de conhecimentos especializados, fiabilidade e recursos – quanto à execução de medidas técnicas e organizativas que cumpram os requisitos do Regulamento. E, por consequência, para evidenciar o cumprimento do responsável pelo tratamento que naquele pretende externalizar certas operações por si realizadas com todas as exigências a que se encontra adstrito ao abrigo do art. 28.º, n.º 1, do RGPD[27].

Finalmente, cumpre ainda salientar a possibilidade de os citados Códigos poderem ser utilizados como um mecanismo legitimador de transferências de dados pessoais para países terceiros ou organizações internacionais[28], constituindo "garantias adequadas", nos termos e para os efeitos previstos no regime especial de proteção, contemplado no Capítulo V do Regulamento[29]. Tudo isto, note-se, sem esquecer que a adesão e cumprimento com um instrumento de tal natureza, por parte de um responsável pelo tratamento, sempre deverá ser considerada como um *fator atenuante*, aquando da eventual decisão da competente autoridade de controlo lhe aplicar ou não uma coima, e, em caso positivo, de determinar o respetivo montante, em caso de violação do RGPD[30].

---

26) No mesmo sentido parece, aliás, apontar o Considerando 77 do RGPD, segundo o qual "[a]s orientações sobre a execução de medidas adequadas e sobre a comprovação de conformidade pelos responsáveis pelo tratamento ou subcontratantes, em especial no que diz respeito à identificação dos riscos relacionados com o tratamento, à sua avaliação em termos de origem, natureza, probabilidade e gravidade, bem como à identificação das melhores práticas para a atenuação dos riscos, poderão ser obtidas nomeadamente recorrendo a Códigos de Conduta aprovados (...)".

27) Recorde-se que, de acordo com o referido preceito "[q]uando o tratamento dos dados for efetuado por sua conta, o responsável pelo tratamento recorre apenas a subcontratantes que apresentem garantias suficientes de execução de medidas técnicas e organizativas adequadas de uma forma que o tratamento satisfaça os requisitos do presente regulamento e assegure a defesa dos direitos do titular dos dados". Para mais desenvolvimentos, veja-se, por todos, o nosso ALVES, Joel A., *Administração Pública e Dados Pessoais: da necessidade do tratamento ao imperativo da proteção*, cit., pp. 309 e ss.

28) Cfr. art. 46.º, n.º 2, al. e) do RGPD.

29) Para uma introdução ao tema, cfr., por todos, ALVES, Joel A., *Administração Pública e Dados Pessoais: da necessidade do tratamento ao imperativo da proteção*, cit., pp. 315 e ss.

30) Assim decorre do art. 83.º, n.º 2, al. j), do RGPD, o qual estabelece que "[*a]o decidir sobre a aplicação de uma coima e sobre o montante da coima em cada caso individual, é tido em devida consideração o seguinte:* (...)

IV. POTENCIAIS VANTAGENS

Do exposto, não parece árduo de compreender que os *Código de Conduta* se apresentam como instrumentos merecedores da maior atenção por parte de responsáveis pelo tratamento e subcontratantes. É que, para além da *certeza* e *segurança* jurídicas que estes proporcionam – desde logo, como consequência das funções de *facilitação* e *demonstração* do cumprimento de que anteriormente tratamos –, muitas outras são as potenciais vantagens que poderão resultar da eventual elaboração e/ou adesão aos mesmos:

a) Em primeiro lugar, porque os Códigos de Conduta fomentam a autonomia de responsáveis pelo tratamento e subcontratantes[31], diminuindo, nomeadamente, a sua dependência face às autoridades de controlo, no sentido de estas emanarem orientações mais exaustivas para as suas atividades de tratamento específicas[32];

b) Em segundo lugar, porque os citados instrumentos incentivam o desenvolvimento de uma abordagem *coletiva* e *coerente* a problemas e necessidades transversais a um determinado setor[33]. Circunstância que favorece não só o *diálogo* entre os seus diferentes membros, como também a própria *consolidação de melhores práticas*[34], tendo em conta as suas experiências no terreno;

c) Em terceiro lugar, porque os Códigos em causa podem ajudar a conquistar a *confiança* dos titulares dos dados: seja por promoverem uma cultura de transparência (nomeadamente, através da possível abordagem de aspetos que tenham sido suscitados por preocupações do público em geral, ou por preocupações existentes no próprio setor)[35]; seja, tão-somente, pela decisão de adesão aos mesmos, por parte de um dado responsável pelo tratamento ou subcontratante, contribuir para a perceção de que este encara o respeito pelos princípios e regras aplicáveis em matéria de proteção de dados pessoais como algo sério[36];

---

j) *o cumprimento de Códigos de Conduta aprovados nos termos do artigo 40.º* ou de procedimento de certificação aprovados nos termos do artigo 42.º" *(destaque nosso)*.

31 Cfr. Comité Europeu para a Proteção de Dados, "Diretrizes 1/2019 relativas aos Códigos de Conduta e aos Organismos de Supervisão ao abrigo do Regulamento (UE) 2016/679", p. 10.

32 Idem, ibidem.

33 Idem, ibidem.

34 Idem, ibidem.

35 Idem, ibidem.

36 Cfr. VOIGT, Paul / VON DEM BUSSCHE, Axel, *The EU General Data Protection Regulation (GDPR): a practical guide,* cit., p. 71.

d) Em quarto e último lugar porque, conforme já foi salientado pelo Comité Europeu da Proteção de Dados, tais Códigos representam uma oportunidade para que os responsáveis pelo tratamento e subcontratante estabeleçam um conjunto de regras que contribuam para a correta aplicação do RGPD "de uma forma potencialmente eficaz em termos de custos"[37] – o que não pode deixar de ser encarado como um fator relevante, em especial, para entidades da Administração Pública atentas as exigências de *economicidade* a que a sua atuação se encontra adstrita, nos termos do art. 5.º do Código do Procedimento Administrativo (*princípio da boa administração*)[38].

V. REQUISITOS DE ADMISSIBILIDADE E APROVAÇÃO

Acontece que, para que todas as referidas potencialidades possam ser aproveitadas, sempre terá o Código em questão de ser submetido à apreciação da autoridade de controlo competente[39] - *in casu*, a Comissão Nacional de Proteção de Dados[40] - nos termos e para os efeitos previstos no art. 40.º, n.º 5, do RGPD[41].

a) Requisitos de admissibilidade

De todo o modo, ponto é que o Comité Europeu para a Proteção de Dados teve já oportunidade de clarificar que a suprarreferida autoridade apenas procederá à avaliação do

37 Idem, ibidem.

38 Sobre este princípio, veja-se o nosso ALVES, Joel A., "Administração eletrónica, eficiência e proteção de dados: breves considerações à luz dos princípios gerais da atividade administrativa", in *Anuário da Proteção de Dados 2022* (coord. Francisco Pereira Coutinho / Graça Canto Moniz), CEDIS, Lisboa, 2022, pp. 136 e ss.

39 Autoridade essa que, no caso de Códigos relativos a atividades de tratamento de dados pessoais realizadas num único Estado-Membro - como será, ao que tudo indica, aqui o caso – será a "autoridade de controlo nacional para os membros a quem o código se aplica". Nesse sentido, cfr. Comité Europeu para a Proteção de Dados, "Diretrizes 1/2019 relativas aos Códigos de Conduta e aos Organismos de Supervisão ao abrigo do Regulamento (UE) 2016/679", cit., p. 18, nota 54.

40 Recorde-se que, nos termos do art. 3.º da Lei n.º 58/2019, de 8 de agosto, a Comissão Nacional de Proteção de Dados qualifica-se como a autoridade de controlo nacional, para efeitos do RGPD.

41 Assim, e de acordo com tal preceito, o Código - ou mais precisamente, o Projeto de Código - submetido à Comissão será sujeito a um procedimento avaliativo, após o que a referida autoridade emitirá um parecer fundamentado, pronunciando-se sobre a sua conformidade com o Regulamento, e, bem assim decidindo-se ou não pela sua aprovação, consoante este apresente ou não garantias apropriadas suficientes. Em caso positivo – e na medida em que se confirme que tal instrumento não se relaciona com atividade de tratamento de dados pessoais realizadas para além do Estado Português –, deverá, finalmente, a Comissão proceder ao seu registo e publicação, nos termos do art. 40.º, n.º 6, do Regulamento. Para mais desenvolvimentos, cfr. Comité Europeu para a Proteção de Dados, "Diretrizes 1/2019 relativas aos Códigos de Conduta e aos Organismos de Supervisão ao abrigo do Regulamento (UE) 2016/679", cit., pp. 18 e ss.

conteúdo do Código propriamente dito caso entenda que tal instrumento satisfaz um conjunto mínimo de *requisitos de admissibilidade*[42].

Assim, importa, antes de tudo o mais, assegurar que o Projeto de Código proposto[43]:

i) É acompanhado de uma exposição de motivos e de toda a documentação de apoio que se possa afigurar relevante para a sua avaliação;

ii) É submetido por uma entidade dotada de legitimidade para a sua elaboração, nos termos do RGPD[44], e que, ademais, fornece informações suscetíveis de demonstrar que é um organismo capaz de compreender as necessidade do setor específico que representa;

iii) Define claramente a atividade ou o setor de tratamento que se propõe regular, bem assim como os problemas que se destina a resolver;

iv) Identifica o seu âmbito de aplicação de territorial (nacional ou transnacional);

v) Fornece informações que justifiquem a qualificação da autoridade de controlo a quem o Código será formalmente apresentado como a autoridade que é efetivamente competente para esse mesmo efeito;

vi) Contempla procedimentos que possibilitem uma supervisão eficaz do cumprimento das suas disposições por parte dos responsáveis pelo tratamento e/ou subcontratantes que se comprometam a aplicá-lo;

vii) Identifica um organismo responsável por realizar essa supervisão, explicando como este observará os requisitos associados a tal tarefa;

viii) Inclui informações relativamente à realização da consulta a que alude o considerando 99 do Regulamento[45], e ao seu respetivo alcance;

ix) Encontra-se em conformidade com a pertinente legislação nacional aplicável;

---

42 Cfr. Comité Europeu para a Proteção de Dados, "Diretrizes 1/2019 relativas aos Códigos de Conduta e aos Organismos de Supervisão ao abrigo do Regulamento (UE) 2016/679", pp. 18-19.

43 Seguimos aqui de perto o estabelecido no documento "Apêndice 3 – Lista de Verificação para a Apresentação", anexo às citadas "Diretrizes 1/2019 relativas aos Códigos de Conduta e aos Organismos de Supervisão ao abrigo do Regulamento (UE) 2016/679" do Comité Europeu para a Proteção de Dados.

44 Recorde-se que, nos termos do art. 40.º, n.º 2, do RGPD, tal legitimidade assiste às "associações e outros organismos representantes de categorias de responsáveis pelo tratamento ou de subcontratantes".

45 Com efeito, dispõe o supramencionado considerando que: "durante o processo de elaboração de um código de conduta, ou na sua alteração ou aditamento, as associações e outros organismos representantes de categorias de responsáveis pelo tratamento ou de subcontratantes deverão consultar as partes interessadas, nomeadamente os titulares dos dados, se possível, e ter em conta os contributos recebidos e as opiniões expressas em resposta a essas consultas".

x) Cumpre com os requisitos linguísticos da autoridade a que será apresentado[46].

b) Requisitos de aprovação

Acresce que, para que o sobredito Projeto possa merecer a aprovação da Comissão Nacional de Proteção de Dados, será ainda imperativo garantir o preenchimento de cinco requisitos adicionais, de ordem material. Em concreto, deve a documentação apresentado à autoridade de controlo evidenciar que o Código:

i) Satisfaz uma necessidade específica do setor e/ou da atividade de tratamento de dados pessoais que se propõe disciplinar, explicando e definindo quais os problemas que aquele se destina a abordar, bem como a forma como as soluções nele disponibilizadas se podem revelar eficazes e benéfica, seja para os seus potenciais aderentes, seja para os próprios titulares dos dados[47];

ii) *Facilita a aplicação efetiva do RGPD*, estipulando com clareza como os seus requisitos devem ser implementados, ao nível do setor em causa, e identificando e suprindo as necessidades do mesmo[48];

iii) *Especifica a aplicação do Regulamento*, não se limitando a reproduzir os princípios e regras nele contidos, mas antes, concretizando-os em função das nuances do setor a que se destinam, conferindo *valor acrescentado* ao ordenamento jurídico vigente[49];

iv) *Oferece garantias adequadas e eficazes para a atenuação dos riscos relacionados as atividades de tratamento que visa disciplinar*[50];

v) *Contempla procedimentos que garantam a supervisão adequada das suas regras*, bem como a aplicação de medidas de execução eficientes e significativas para garantir o pleno cumprimento dessas regras[51].

46 Note-se que, de acordo com o art. 54.º do Código do Procedimento Administrativo, "a língua do procedimento é a língua portuguesa"; pelo que, entre nós, será esta a língua em que qualquer Projeto de Código deve ser submetido à apreciação da Comissão Nacional de Proteção de Dados.
47 Cfr. Comité Europeu para a Proteção de Dados, "Diretrizes 1/2019 relativas aos Códigos de Conduta e aos Organismos de Supervisão ao abrigo do Regulamento (UE) 2016/679", p. 15.
48 Idem, p. 16.
49 Idem, pp. 16-17.
50 Idem, p. 17.
51 Idem, p. 18.

VI. CONTEÚDO

Já no que concerne aos domínios suscetíveis de serem concretamente regulados por um tal instrumento, cumpre recordar que o RGPD se limita a estabelecer, no seu art. 40.º, n.º 2, um *elenco exemplificativo* de matérias onde a aplicação do Regulamento pode ser especificada. Daí que, ao menos em tese, um *Código de Conduta* tanto possa versar sobre *todos* os aspetos constantes desse mesmo elenco, apenas sobre *alguns* deles, ou ainda sobre quaisquer *outros* que sejam disciplinados pelo RGPD.

Seja como for, nem por isso o sobredito preceito deve deixar de ser encarado como um importante referencial: não só por fornecer pistas concretas quanto aos domínios que poderão ser abordados ao nível de um tal Código; como também, por refletir muitas das dimensões jurídicas que poderão conferir um efetivo *valor acrescentado* a um instrumento de tal recorte[52] - requisito que, como vimos, se demonstra essencial para que este venha a ser aprovado pela competente autoridade nacional de controlo, nos termos do art. 40.º, n.º 5, do Regulamento.

VII. NOTAS FINAIS

De acordo com a informação publicamente disponibilizada pelo Comité Europeu para a Proteção de Dados, foram, até à presente data aprovados, cerca de trinta e um Códigos de Conduta[53].

No que especificamente respeita à realidade portuguesa, contam-se única e exclusivamente dois: (i) o Regulamento Intermunicipal de Dados Pessoais da Comunidade Intermunicipal do Tâmega e Sousa (aprovado em 19 de janeiro de 2022); e (ii) o Código de Conduta para o Tratamento de Dados Pessoais da APDL - Administração dos Portos do Douro, Leixões e Viana do Castelo, S.A. (aprovado em 22 de março de 2022).

52Nesse sentido, sustentando que o *valor acrescentado* de um Código de Conduta dificilmente poderá ser avaliado pela competente autoridade de controlo, caso o mesmo não aborde qualquer das matérias enunciadas no art. 40.º, n.º 2, do RGPD, cfr. GÓMEZ, Alberto Díaz-Romeral, "Los códigos de conducta en el Reglamento General de Protección de Datos", cit., p. 400.

53Recorde-se que, nos termos do art. 40.º, n.º 11, do RGPD, "o Comité recolhe todos os códigos de conduta aprovados, respetivas alterações e respetivos aditamentos num registo e disponibiliza-os ao público pelos meios adequados". O registo em questão encontra-se disponível em https://www.edpb.europa.eu/registers/register-of-consistency-and-of-accountability-tools/codes-of-conduct_pt [Consultado a 29 de setembro de 2026].

Nesta senda, não podemos deixar de manifestar alguma surpresa - e diríamos até, perplexidade - face ao aparente desinteresse que a generalidade dos responsáveis pelo tratamento e subcontratantes tem demonstrado em relação a estes instrumentos. Circunstância que, quanto a nós, se tende a dever mais ao desconhecimento sobre as reais potencialidades dos Códigos de Conduta do que propriamente à falta de crença nos mesmos. Com efeito, importa que as autoridades de controlo nacionais possam ajudar a inverter este paradigma, promovendo uma maior consciencialização sobre as vantagens jurídicas (e não só) dos supramencionados Códigos e, bem assim, *incentivando a sua elaboração* no respetivo território, tal como é sua atribuição nos termos do art. 57.º, n.º 1, al. m), do RGPD[54]. Sem embargo, cumpre não esquecer que, em Portugal, o art. 15.º, n.º 2, da Lei n.º 58/2019, de 8 de agosto, veio expressamente determinar que "*o tratamento de dados pessoais pela administração direta e indireta do Estado é objeto de códigos de conduta próprios*". Como tal, parece-nos inevitável que também o Governo, ao abrigo das competências administrativas que lhe são constitucionalmente conferidas, tenha uma palavra a dizer nesta matéria - pelo menos, no que especificamente respeita a esta parte do setor público[55].

54Conforme explicitamente se lê neste preceito, "sem prejuízo de outras atribuições previstas nos termos do presente regulamento, cada autoridade de controlo, no território respetivo: (...) *incentiva a elaboração de códigos de conduta nos termos do artigo 40.º, n.º 1* (...)" (destaque nosso). Concretizando esta disposição, o art. 15.º, n.º 1, da Lei n.º 58/2019, de 8 de agosto, estabelece, por sua vez, entre nós, que "*compete à CNPD fomentar a elaboração de códigos de conduta que regulem atividades determinadas* (...) (destaque nosso).

55Assim é, pois que, nos termos do art. 199.º, n.º d), da Constituição da República Portuguesa, "compete ao governo "*dirigir os serviços e a atividade da administração direta do Estado*, civil e militar, *superintender na administração indireta e exercer a tutela sobre esta* e sobre a administração autónoma" (destaque nosso).



FROM COMPLIANCE TO INTEGRABILITY
DATA PROTECTION IN M&A TRANSACTIONS

DIOGO GUERREIRO DUARTE *

* Master's degree in International Law, Faculty of Law, University of Lisbon.

RESUMO:

As operações de fusão e aquisição podem gerar riscos em matéria de proteção de dados pessoais que vão além das questões de conformidade histórica normalmente analisadas num processo convencional de *due diligence*. O Regulamento Geral sobre a Proteção de Dados (RGPD) não estabelece um regime específico para estas operações. Ainda assim, uma aquisição ou fusão pode alterar quem tem acesso aos dados pessoais, a repartição de responsabilidades pelo tratamento, as finalidades prosseguidas, os sistemas utilizados, os destinatários dos dados e os respetivos fluxos, incluindo no contexto de transferências internacionais. O presente artigo analisa estas alterações a partir do ciclo de vida da transação e distingue três categorias de risco: risco da transação, risco preexistente e risco de integração. Sustenta-se que cada operação de tratamento materialmente relevante, bem como cada transição entre diferentes formas de tratamento, deve ser apreciada autonomamente, uma vez que a conclusão de uma transação não amplia, por si só, a licitude do tratamento dos dados pessoais associados ao negócio adquirido.

Partindo desta distinção, o artigo desenvolve uma metodologia específica para o adquirente, designada *Data Protection Integration Readiness*, destinada a avaliar se, e em que condições, o tratamento de dados pessoais associado à sociedade-alvo pode continuar após a aquisição, ser migrado para outros sistemas, ser combinado com outros conjuntos de dados, ser utilizado para novas finalidades ou, de outro modo, ser integrado no modelo operacional pretendido pelo adquirente. A análise transpõe para o contexto das fusões e aquisições princípios e obrigações já previstos no direito da União Europeia em matéria de proteção de dados, designadamente os relativos à licitude do tratamento, à limitação das finalidades, à determinação das funções e responsabilidades dos intervenientes no tratamento, ao dever de assegurar e demonstrar a conformidade, à segurança, aos direitos dos titulares dos dados, às relações com subcontratantes e às transferências internacionais de dados. Por fim, o artigo propõe um modelo assente em cinco pontos de decisão, através dos quais a análise em matéria de proteção de dados pode orientar a divulgação e transmissão de dados, a gestão dos riscos identificados, a continuidade das operações de tratamento e a integração pós-aquisição, assegurando que as questões jurídicas relevantes são apreciadas antes de ocorrerem alterações materialmente relevantes no tratamento.

Palavras-chave: fusões e aquisições; proteção de dados; Regulamento Geral sobre a Proteção de Dados; due diligence; risco de integração; Data Protection Integration Readiness; integração pós-aquisição.

ABSTRACT:

Mergers and acquisitions create data protection risks that go beyond the historical compliance issues usually reviewed through conventional due diligence. The General Data Protection Regulation does not establish a transaction-specific regime, yet an acquisition may alter access to personal data, Controller relationships, processing purposes, systems, recipients and entail international data flows. This article examines those changes through a lifecycle approach that distinguishes three categories of risk: transaction risk, inherited risk and integration risk. It argues that each material processing operation or transition should be assessed on its own terms and that closing does not, by itself, enlarge the lawful uses of personal data associated with the acquired business.

The article develops an acquirer-specific methodology, Data Protection Integration Readiness, for assessing whether, and under what conditions, the target's processing may lawfully continue, be migrated, combined, repurposed or otherwise incorporated into the intended post-closing operating model. The analysis applies existing European Union data protection principles on lawful basis, purpose limitation, Controllership, accountability, security, data subject rights, Processor governance and international transfers to the transactional lifecycle. It concludes with a five-gate framework through which data protection analysis can inform disclosure, risk treatment, processing continuity and post-closing integration before material processing transitions take place.

Keywords: Mergers & Acquisitions; data protection; General Data Protection Regulation; due diligence; integration risk; Data Protection Integration Readiness; post-closing integration.

I. INTRODUCTION

When Marriott International acquired Starwood in 2016, it acquired systems that had already been compromised two years earlier. The intrusion remained undetected after the acquisition and continued until 2018. The significance of this case for data protection is narrower than a general warning about inadequate diligence, and its importance does not lie in treating the business acquisition as a transmission of Starwood's historical liability to Marriott, nor in assuming that more extensive transaction diligence would necessarily have revealed the compromise. The Information Commissioner itself expressly declined to make such a finding. What became relevant was Marriott's own conduct after the General Data Protection Regulation (GDPR) became applicable, namely, its own continued processing of personal data through the acquired systems and its responsibility for guaranteeing that adequate security measures were implemented.[1] The episode exposes a problem that goes beyond cybersecurity. The transactional problem made evident by this case is that an acquisition may change access to personal data, Controller relationships, systems, purposes and intended uses without those changes being resolved by the corporate transaction itself.

The GDPR contains no special regime for mergers and acquisitions (M&A), and its obligations follow the processing, regardless of the labels attached to the different stages of a transaction. M&A nevertheless creates identifiable changes in processing that alter the legal questions the parties must answer. Disclosure during diligence, changes in access or Controllership at completion, continuance of existing processing, migration between systems, combination of datasets and introduction of new purposes for the processing may each constitute materially different processing operations or transitions. These events do not necessarily form a single processing operation. Each material operation, or related set of operations, must be assessed by reference to the Controller or Controllers involved, the purposes pursued, the applicable lawful basis and any other relevant conditions.

1 See Information Commissioner's Office, *Marriott International, Inc., Monetary Penalty Notice*, 30 October 2020. Marriott stated that it had been able to conduct only limited due diligence on Starwood's data processing systems and databases during the acquisition. The Commissioner expressly declined to determine whether more extensive due diligence was possible during the takeover and limited the infringement findings to Marriott's conduct after the GDPR became applicable. The decision found that Marriott continued processing through acquired Starwood systems without appropriate security measures. See especially paragraphs 3.3-3.4 and 6.62-6.68.

Historical compliance remains an essential part of that assessment, but it captures only one aspect of the transaction's data protection risk. In this article, I argue that data protection risk in M&A should not be treated as a single due diligence exercise focused principally on the target's historical compliance. Historical compliance continues to be essential, but it captures only one part of the transaction's data protection problem. The article distinguishes three categories of risk: (i) transaction risk, which results from the processing undertaken to evaluate, negotiate, execute and give effect to the transaction; (ii) inherited risk, which covers pre-existing liabilities, vulnerabilities, constraints and exposures whose consequences remain relevant after acquisition; and (iii) integration risk, which concerns the future state of the processing by way of addressing the question of knowing whether the target's processing can lawfully continue, migrate, combine, be repurposed or otherwise form part of the acquirer's intended processing environment.

From this taxonomy follows the article's central organising proposition, which emphasises that closing does not, by itself, expand the lawful uses and purposes of the processing of personal data associated with the acquired business. Corporate authority over a company or a set of assets is distinct from legal authority to process the personal data connected with them. Even a favourable assessment of the target's historical compliance does not establish that the acquirer may lawfully implement its intended post-closing uses. Corporate acquisition can confer access to, or at least, put personal data within the acquirer’s reach, without conferring equivalent rights of use or further processing.

This distinction and the gap it contain lead to a second inquiry that conventional compliance diligence does not necessarily answer. The article develops the concept of Data Protection Integration Readiness to assess whether, and under what conditions, the target's processing operations can lawfully form part of the acquirer's intended future state. This concept creates no new obligation under the GDPR, but it works as an acquirer-specific methodology for applying existing obligations prospectively to processing transitions that occur during M&A processes. Data protection maturity (or “Privacy Maturity,” as it is often described) can be used to describe the target, whereas Data Protection Integration Readiness describes the relationship between the target, the acquirer, and the processing environment the acquirer intends to create.

The proposition that data protection work goes past pre-transaction diligence is not new. Existing scholarship and practitioner guidance have already examined data protection, information security, and information governance risks across the M&A lifecycle, including during post-closing integration, and have recognised that due diligence may need to consider intended post-acquisition uses of personal data and its feasibility, particularly where the acquirer intends to combine, exploit or otherwise use the target's data after completion.[2] This article builds on that literature but addresses a more specific question that remains underdeveloped: how to distinguish the target's existing compliance from its legal capacity to enter the processing environment contemplated by a particular acquirer. The concept of Data Protection Integration Readiness developed in this article proposes to address that question by treating each material processing transition as the relevant unit of legal analysis, rejecting the existing legal-practice idea that the corporate transaction constitutes a single data protection event.

The individual legal principles assembled in this analysis are themselves already established. The distinction between share and asset transactions, the treatment of disclosure during diligence and at completion, functional Controllership, purpose limitation, security obligations and the legal significance of dataset combination are already reflected in legislation, case law, supervisory guidance and transactional practice.[3] The German supervisory guidance, for example, distinguishes a share acquisition, which typically does not itself entail the transmission of personal data to the acquirer, from an asset transaction,

2 See James A. Sherer, Taylor M. Hoffman and Eugenio E. Ortiz, "Merger and Acquisition Due Diligence: A Proposed Framework to Incorporate Data Privacy, Information Security, E-Discovery, and Information Governance into Due Diligence Practices" (2015) 21 *Richmond Journal of Law & Technology* 5; Matthew Knouff, Rani Habash and Brad Janssen, "Data Privacy and Security Issues in M&A Transactions: Part One", International Association of Privacy Professionals, 26 April 2016, and "Data Privacy and Security Issues in M&A Transactions: Part Two", International Association of Privacy Professionals, 22 August 2016; The Sedona Conference, "Commentary on Data Privacy and Security Issues in Mergers & Acquisitions Practice" (2019) 20 *Sedona Conference Journal* 233; Shane Woulfe, *GDPR Considerations in M&A Transactions: How Can the Risk of Non-Compliance with the GDPR Be Reduced in the M&A Process from a Buyers' Perspective?* (LL.M. thesis, Tilburg University 2020); Justin Weiss and Marcin Czarnecki, *Privacy in M&A Transactions Playbook* (International Association of Privacy Professionals, 2 December 2021). Most of these materials recognise that data protection and information security considerations may extend throughout the transaction lifecycle and into post-closing integration, and that transaction diligence may need to take account of intended future uses and integration requirements.

3 See, among others, Datenschutzkonferenz, *Übermittlungen personenbezogener Daten an die Erwerberin oder den Erwerber eines Unternehmens im Rahmen eines Asset-Deals* ("DSK, *Asset-Deals"*), Decision of 11 September 2024; European Data Protection Board, *Statement on Privacy Implications of Mergers* ("EDPB, *Statement on Mergers"*), adopted 19 February 2020.

where often disclosure and subsequent processing of personal data must be examined according to the nature of data and circumstances involved. The European Data Protection Board (EDPB) has likewise drawn attention, in the context of the Google/Fitbit transaction, to the risks associated with the further combination and accumulation of personal data following an acquisition.[4] The contribution of this article is therefore one of synthesis and methodology as it reorganises these materials and the existing principles around the processing transitions generated by an acquisition. The contribution of this article also lies in bringing these strands together within a legal methodology that connects the target's present condition to the feasibility of the acquirer's intended future processing environment. In particular, the three-part taxonomy identifies where risk arises; the Data Protection Integration Readiness tests whether the intended integration is legally feasible; and the transition-specific analysis determines when the relevant legal question must be answered.

A boundary should also be drawn, and this article does not ignore it. There is a substantial and distinct body of work concerning data protection in merger control, including whether the accumulation or combination of personal data should influence competition-law assessment of a concentration.[5] That is not the principal subject of this article. The question examined here is transactional and aims to analyse how data protection risk arises and changes as the parties evaluate, complete and integrate an acquisition. The two inquiries may overlap, especially where a transaction contemplates combining datasets, but they ask different legal questions.

The argument develops through six substantive sections. Section 2 sets out the lifecycle analysis and the three-part taxonomy of transaction, inherited and integration risk. Section 3 examines transaction risk before closing, with particular attention to the lawfulness of disclosure and review during due diligence. Data Protection Integration Readiness is developed in Section 4 through seven dimensions used to assess the intended future state.

4 See EDPB, *Statement on Mergers* (footnote n 3).

5 See, among others, European Data Protection Supervisor, *Privacy and Competitiveness in the Age of Big Data: The Interplay between Data Protection, Competition Law and Consumer Protection in the Digital Economy*, Preliminary Opinion, 26 March 2014; Ben Holles de Peyer, "EU Merger Control and Big Data" (2017) 13(4) *Journal of Competition Law & Economics* 767-790; Maria C. Wasastjerna, "The Role of Big Data and Digital Privacy in Merger Review" (2018) 14(2-3) *European Competition Journal* 417-444; and Robin Vandendriessche and Caroline Buts, "Data Protection Considerations in Competition Law Assessments: A Qualitative Document Analysis of EU Decision Texts" (2025) 21(1) *Journal of Competition Law & Economics* 1-43. These materials show the distinct debate concerning whether and how personal data accumulation, data combination and data protection considerations should enter competition law and merger control analysis.

Section 5 addresses risk treatment at signing and closing, including the limits of contractual allocation and the need to plan for processing continuity. Section 6 considers the legal constraints that arise during post-closing transformation and integration. Section 7 brings the analysis together in a five-gate framework under which the relevant legal determinations precede any material transition in processing. Section 8 concludes.

II. DATA PROTECTION RISK ACROSS THE M&A LIFECYCLE: A THREE-PART TAXONOMY

In mergers and acquisitions, data protection, as a field of law, is typically treated as a separate workstream, often labelled "privacy due diligence", and evaluated in parallel with other legal domains such as employment, tax, contract, intellectual property, and environmental law. The scope and materiality thresholds of these segmented but complementary workstreams are established to review the target's documents and practices, identify any relevant risks, assess their materiality and inform how they should affect valuation, contractual protection, remediation or other transaction decisions. Findings from each legal field are then consolidated, often into a single due diligence report, to present an overall view of the target's compliance and the consequences of identified risks for the transaction.[6] Although this process is a necessary initial step, it is not sufficient on its own. Historical compliance reveals the target's existing risk profile, but it looks mainly backwards. If it becomes the sole frame of analysis, legally distinct processing operations may be compressed into one undifferentiated inquiry. It is necessary to consider that significant risk also arises from the acquirer's intended future processing activities.

One common analytical mistake is the assumption that acquiring a business necessarily transmits its personal data to a new Controller, generally the acquirer. Whether such a transmission occurs depends on the structure of the transaction itself. In a share acquisition, the acquirer acquires an ownership interest in the target, which generally continues as the same legal entity. If the target retains control over the purposes and means of processing

6 See FTI Technology, *Data Risks and Challenges in M&A Transactions*; Bird & Bird, *M&A Post Completion - Data Protection Integration and Gap Closing*. These practitioner materials present the treatment of data protection as a distinct transactional workstream and the use of diligence findings to inform risk allocation, remediation and post-completion integration.

personal data, its status as Controller remains unchanged despite a change in ownership. By contrast, an asset acquisition may involve the disclosure of customer, employee, supplier or other personal data to a different entity, after which some of that data may be processed by the acquirer as (the new) Controller. The lawfulness of that sequence cannot be assessed in the abstract. It depends on the personal data disclosed, the processing contemplated and the circumstances of the transaction.[7]

Corporate control and Controllership under the GDPR are separate concepts. Article 4(7) GDPR defines the Controller as the entity that determines the purposes and means of processing, and not by reference to ownership of the entity carrying out that processing. The EDPB adopts the same functional approach, directing attention to what the parties actually do instead of their formal designations.[8] The Court of Justice of the European Union (CJEU) has affirmed this functional approach in Nacionalinis visuomenės sveikatos centras. It held that an entity may qualify as a Controller if it has participated in determining the purposes and means of processing, even though it did not perform the processing itself.[9] Therefore, Controllership depends on the determination of purposes and means on a functional basis, not on ownership, and, for the present analysis, acquiring a target does not, by itself, make the acquirer the Controller of the target's processing activities.

This article introduces a taxonomy of three types of risks that follow the transaction lifecycle: (i) transaction risk, (ii) inherited risk, and (iii) integration risk. Its premise is that the transaction lifecycle should be understood as a sequence of distinct processing operations, not simply as a series of corporate or deal stages. This distinction is significant because the GDPR, particularly under Article 5, applies to the processing operations in question regardless of whether the transaction describes them as "diligence", "closing" or "integration". What changes throughout the transaction is not the law, but the specific

7 See DSK, *Asset-Deals* (footnote n 3), pp 1-2. The German Data Protection Conference distinguishes share transactions, in which the legal entity ordinarily continues unchanged and, apart from processing undertaken during due diligence, no transmission of personal data generally occurs merely through the change in ownership, from asset transactions, which require a differentiated assessment of disclosure and subsequent processing.

8 See European Data Protection Board, *Guidelines 07/2020 on the Concepts of Controller and Processor in the GDPR* ("EDPB, *Guidelines 07/2020"*), Version 2.1, adopted 7 July 2021, paras 12-15. In these Guidelines, EDPB refers to the Controller and Processor as functional concepts determined by the parties' actual roles and decision-making authority instead of their formal designation or corporate ownership.

9 See Case C-683/21, *Nacionalinis visuomenės sveikatos centras prie Sveikatos apsaugos ministerijos v Valstybinė duomenų apsaugos inspekcija*, judgment of 5 December 2023, paras 28-37. The Court confirmed that an entity may qualify as Controller where it participates in determining the purposes and means of processing even if it does not itself carry out the processing operation.

operations to which it applies. Each material operation or set of operations must therefore be assessed by reference to the relevant Controller(s), purposes of the processing, lawful basis and other applicable legal requirements and conditions. Where further processing for another purpose falls within Article 6(4) GDPR, its compatibility with the original purpose must also be assessed.[10] For example, a single customer dataset may undergo three successive processing operations: (1) review in the data room by or on behalf of the acquirer, (2) disclosure to the acquirer in connection with the closing of an asset transaction, and (3) integration into the acquirer's customer database for cross-selling. Each one of these operations raises a separate question of lawfulness, and describing the whole sequence as "M&A processing" conceals those differences. A lifecycle model is particularly useful in this domain because it identifies when a transaction changes factual circumstances relevant to the GDPR, without assigning legal consequences solely by reference to the deal's stage.

Transaction risk covers the processing activities that take place to evaluate, negotiate, execute and give effect to the transaction. It includes the disclosure of personal data via a virtual data room, access for bidders and advisers, clean-team arrangements, management presentations, and cross-border review. These activities exist because the transaction is taking place, and, therefore, their lawfulness must be assessed separately from the lawfulness of the target's underlying business operations.

Inherited risk concerns pre-existing liabilities, vulnerabilities, restrictions and exposures whose legal, economic or operational consequences remain relevant after the acquisition. Examples include unlawful processing, excessive retention of personal data, unresolved personal data breaches, pending complaints or investigations and inadequate security. Historical compliance findings are evidence of inherited risk, but the two concepts are not interchangeable. A past breach that has been properly remediated may leave little residual exposure. An unresolved vulnerability may represent a serious inherited risk even if it has not yet resulted in a complaint, investigation or enforcement action, and may be missed by a review focused mainly on documents.[11] The term "inherited" is used analytically in this

10 See Regulation (EU) 2016/679, Articles 4(2), 6(1) and 6(4), recital 50; EDPB, *Guidelines 07/2020* (footnote n 8), para 42.

11 See Information Commissioner's Office, *Marriott International, Inc., Monetary Penalty Notice* (footnote n 1); see also Section 4. The Marriott case is used here to show how an unresolved vulnerability originating before acquisition may remain legally and operationally relevant afterwards.

context and does not suggest that the target's historical liability under the GDPR automatically passes to the acquirer. Conventional data protection due diligence, nevertheless, has a central role in identifying and characterising this category of risk.

Integration risk is different in both timing and focus. It may not have materialised at signing or closing, although the acquirer's plans may allow it to prospectively identify them by reference to the acquirer's intended use of the acquired business. The question addressed by this type of risk is whether the target's existing processing may lawfully continue, migrate, be combined, repurposed, or otherwise enter the future processing environment envisaged by the acquirer. Conventional due diligence is least well equipped to answer this question because its principal object is the target as it exists before integration. Integration risk is therefore the focus of this article.

Transaction structure functions as a variable within this taxonomy. Share acquisitions, asset acquisitions, statutory mergers and carve-outs may engage the same taxonomy in different ways because they produce different changes in Controller identity, data flows and contractual relationships. Structure does not alter the taxonomy itself, but it affects how each category operates. It is particularly important to transaction and inherited risk because they produce different changes in Controller identity, data flows and contractual relationships. Structure may determine whether personal data are disclosed at closing, whether Controller identity changes, which contracts remain in force and which legal entity remains responsible for pre-existing liabilities. Integration risk is also structure-sensitive, although its central inquiry remains the legality of the acquirer's intended future processing. Similar intended uses of the personal data by the acquirer may encounter the same substantive restriction even when the legal route to integration differs. Combining two customer datasets, for example, may engage purpose limitation in both a share acquisition and an asset acquisition. What may differ is the identity of the Controller, the route by which the data become available and the applicable transparency or contractual requirements. This example shows that transaction structure changes how this taxonomy operates without treating it as a new category of risk.

Another aspect of the lifecycle is that acquisition may change or amplify the economic significance of pre-existing exposure. Article 83 GDPR uses the concept of an “undertaking” when setting turnover-based limits for administrative fines. In Deutsche Wohnen, the CJEU held that, where the addressee of an administrative fine is or forms part of an undertaking

within the meaning of Articles 101 and 102 of the Treaty on the Functioning of the European Union (TFEU), that concept must be used when calculating the applicable maximum fine under Article 83(4) to (6) GDPR.[12]

The GDPR itself contains no general rule of successor liability for M&A transactions. Where the Controller responsible for an infringement becomes part of a much larger undertaking, however, the undertaking's turnover may increase the economic significance of the enforcement exposure. In ILVA, the Court further clarified that the turnover of the wider economic unit is relevant when the amount of the fine is determined, so that the penalty is effective, proportionate and dissuasive. This does not mean that an acquisition transfers a target's historical GDPR liability to the acquirer, nor that a fine is to be calculated by mechanically applying a percentage to group turnover. Whether historical liability reaches the acquirer depends, among other matters, on the continued legal existence of the target entity, the structure of the transaction and the applicable rules of corporate succession.[13]

The following sections apply this taxonomy to the material processing operations and transitions that arise across the transaction, starting with those that occur before closing.

III. BEFORE CLOSING: TRANSACTION RISK IN DUE DILIGENCE

Evaluating a transaction usually requires information about the target to be exchanged. Where that information includes personal data, such as employee records, customer lists, contracts naming individuals or litigation files, both its disclosure and its review are processing operations. A common error in transactional practice is to treat this exchange as routine, with confidentiality agreements regarded as the principal safeguard. Confidentiality

12 See Case C-807/21, *Deutsche Wohnen SE v Staatsanwaltschaft Berlin*, judgment of 5 December 2023, paras 53-59. The Court distinguished the rules determining responsibility for an infringement from the competition-law concept of an "undertaking", which becomes relevant in the specific context of calculating administrative fines under Article 83(4) to (6) GDPR. Where the addressee of the fine is or forms part of an undertaking, the applicable maximum is calculated by reference to the worldwide annual turnover of that undertaking. See further Section 5 on the distinction between responsibility for an infringement and calculation of the resulting fine.

13 See Case C-383/23, *ILVA A/S*, judgment of 13 February 2025, paras 23-36. The Court confirmed that, where the Controller forms part of an undertaking, the economic capacity represented by that undertaking is relevant not only to the statutory maximum under Article 83 GDPR but also, where appropriate, to determining an effective, proportionate and dissuasive fine. This does not establish a general rule that historical GDPR liability transfers to an acquirer upon acquisition. The attribution of responsibility and the calculation of a fine remain distinct questions.

does not answer the separate question of whether the processing is lawful. Article 6(1)(f) GDPR may frequently be considered as a lawful basis for these processing operations, including the target's interest in pursuing the transaction and the acquirer's interest in assessing it and understanding what it proposes to acquire. Even so, each Controller must nevertheless identify and justify the lawful basis for its own processing.[14] Commercial interests can constitute legitimate interests, provided that the interest pursued is lawful, but this satisfies only the first limb of the Article 6(1)(f) GDPR test. The processing must also be necessary, and the Controller must balance its interests against the rights and interests of the individuals whose data are disclosed or otherwise processed.[15] Neither the Article 29 Working Party's Opinion 06/2014 nor the European Data Protection Board's Guidelines 1/2024 on processing of personal data on Article 6(1)(f) GDPR identifies M&A due diligence as an example of processing grounded on legitimate interests. Reliance on Article 6(1)(f) GDPR must therefore be justified by applying its general requirements to the particular disclosure and review, as there is no transaction-specific regulatory endorsement on which the parties can rely for the processing of personal data.[16]

Routine or comprehensive disclosure cannot be assumed to satisfy the necessity requirement only because it occurs in the context of a lawful and legitimate business operation. Under Article 6(1)(f) GDPR, necessity is a concept distinct from convenience or customary practice. Necessity requires the processing of personal data to be genuinely required for the stated purpose, where the objective of the processing cannot be reasonably achievable through equally effective, less intrusive means.[17] The contents of a data room should therefore be tied

14 On legitimate interests as a possible basis for diligence-stage processing, see GDPR Article 6(1)(f) and recital 47; Article 29 Data Protection Working Party, Opinion 06/2014 on the notion of legitimate interests of the data Controller under Article 7 of Directive 95/46/EC, WP217, 9 April 2014; and European Data Protection Board, Guidelines 1/2024 on processing of personal data under Article 6(1)(f) GDPR, Version 1.0, adopted 8 October 2024. On the possibility that a commercial interest may be a legitimate interest, see Case C-621/22, *Koninklijke Nederlandse Lawn Tennisbond v Autoriteit Persoonsgegevens*, judgment of 4 October 2024, paras 47-50 and 57. The Court confirmed that a commercial interest may qualify as legitimate, but the processing must be strictly necessary and must satisfy the balancing test.

15 On the cumulative character of the three-part test, legitimate interest, necessity, and balancing, WP29, *Opinion 06/2014* (footnote n 14); EDPB, draft *Guidelines 1/2024* (n 14); Case C-252/21, *Meta Platforms Inc and Others v Bundeskartellamt*, judgment of 4 July 2023, ECLI:EU:C:2023:537; *Koninklijke Nederlandse Lawn Tennisbond* (footnote n 14).

16 Note that neither WP29, *Opinion 06/2014* (footnote n 14), nor EDPB, draft *Guidelines 1/2024* (footnote n 14), identifies M&A due diligence as a worked example of processing based on legitimate interests.

17 On the strict interpretation of necessity, under which processing must be genuinely required and the objective must not be reasonably achievable through less intrusive means, see *Meta Platforms* (footnote n 15); *Koninklijke Nederlandse Lawn Tennisbond* (footnote n 14).

to the questions that diligence is intended to answer. An acquirer and its advisers may need to examine the target's data protection practices, the nature and sensitivity of its data holdings, its history of litigation and personal data breaches, or the strength of its consents and contractual arrangements. Those inquiries do not invariably require access to the underlying identifiable data. Aggregate data, samples, redacted documents, and descriptive summaries may provide the reviewer with what they need without exposing individual records. The data minimisation principle set under Article 5(1)(c) GDPR is therefore an operative limit on disclosure that should be considered before personal data is uploaded into the data room and throughout the M&A lifecycle.

Several standard techniques are employed in practice, though they are commonly justified on grounds of confidentiality or competition rather than data protection. A clean team of external advisers may review sensitive material and report its conclusions without giving the bidder access to the underlying data; aggregation, redaction and pseudonymisation can limit the information disclosed to what the diligence question requires; and staged disclosure can postpone access to individual-level data until the transaction is more certain and the need for that access can be shown more clearly. These techniques do not all have the same effect under the GDPR. Aggregation and redaction may reduce the amount of personal data disclosed or, if sufficiently robust, remove personal data from the disclosure altogether. Pseudonymised data remain personal data where attribution remains possible through additional information, although pseudonymisation can reduce the risks to individuals. Clean-team arrangements and access restrictions operate mainly as safeguards relevant to necessity, balancing and security. Viewing these measures through data protection law helps to distinguish those that are legally required from those that are merely prudent.

Where a diligence question can be answered without processing personal data, the most precise application of the data minimisation principle is to keep those data out of the data room. Doing so avoids creating additional processing of personal data by the recipient for that diligence activity. It does not, however, determine whether the acquirer or its advisers fall within the scope of the GDPR in relation to other processing operations.[18]

18 See Articles 4(5) and 5(1)(c), and recital 26 GDPR. Pseudonymisation reduces identifiability and may serve as a safeguard, but it does not take the resulting data outside the scope of the GDPR where attribution remains possible through additional information. See also Case C-413/23 P, *European Data Protection Supervisor v Single Resolution Board*, judgment of 4 September 2025. The Court held that pseudonymised data are not

Special categories of personal data, as defined in Article 9(1) GDPR, require a further layer of analysis. Employee health information, data revealing trade-union membership and other special categories of personal data cannot be processed on the basis of legitimate interests alone. An Article 6 GDPR lawful basis must be accompanied by an applicable exemption under Article 9(2) GDPR. If no such exemption applies to the proposed diligence processing, personal data should not be disclosed in identifiable form. It should generally be withheld, aggregated to a level at which no individual can reasonably be identified, or heavily redacted before being placed in the data room. On the other hand, personal data concerning criminal convictions and offences is subject to an additional restriction under Article 10 GDPR, which allows the processing of such data only under the control of an official authority or where authorised by Union or Member State law, subject to appropriate safeguards. This restriction is easily missed when litigation and investigation files are disclosed.

Several aspects of disclosure are often treated as matters of transactional procedure rather than as legal matters. A data room is itself a processing environment. Article 32 GDPR requirements therefore apply here, requiring Controllers and Processors involved in its operation to adopt technical and organisational measures appropriate to the risk. Access controls, activity logs and confidentiality undertakings are not simply commercial precautions. They may form part of the measures required to secure the processing environment and ensure that personal data stays adequately protected. Where a third party provides the platform for the data room and, in that context, processes personal data on behalf of one of the parties, it must also be engaged on terms that comply with Article 28 GDPR.

The allocation of obligations under the GDPR depends on the roles defined in Articles 4(7) and 4(8) GDPR. Where the target establishes the data room and decides which personal data will be disclosed, who may access them and the transactional purpose for which access is granted, the target will often be the Controller for that disclosure. A virtual data room provider that merely hosts the environment in accordance with the target's documented instructions will likely be a Processor and must be engaged in accordance with Article 28

---

necessarily personal data in the hands of every recipient. Where pseudonymisation effectively prevents a particular recipient from identifying the data subjects, the data may no longer constitute personal data from that recipient's perspective. Identifiability must therefore be assessed in the circumstances of the particular processing operation and from the perspective relevant to the obligation in question. Regarding restrictive pre-completion disclosure in the transaction context, see also DSK, *Asset-Deals* (footnote n 3), section I.

GDPR. External legal advisers do not become Processors for the acquirer or the target merely by accessing its data room. Their status must be assessed by reference to the processing activity they actually perform. A law firm that exercises independent professional judgment when advising its client may be an independent Controller for that processing. An entity engaged only to process documents on another party's documented instructions may instead be a Processor.[19] The relevant classification, as argued by EDPB, is functional and may differ between separate processing activities performed by the same organisations.

The analysis does not change simply because the acquirer or its advisers provide the technical data room environment. Ownership, licensing, or technical operation of the platform is not, by itself, a decisive factor in determining who the Controller is. The seller may operate as a Controller for the disclosure if it decides which personal data are extracted from the target's records, why they are disclosed and who receives access. An entity that determines how the data room will be hosted and secured may assume responsibility for those aspects of the processing, while a third-party hosting provider may act as a Processor for the party instructing it. A law firm may likewise act as an independent Controller when providing legal advice and as a Processor when supplying a separate hosting or document management service, provided that the conditions for each classification are met.

Joint Controllership under Article 26 GDPR requires joint participation in determining the purposes and essential means of the same processing operation. Cooperation in the transaction, or participation in the same information flow, is not enough.[20] Responsibility for data room security and for the Processor chain must accordingly be allocated by reference to what each party actually does in relation to the processing activity.

The Controller's obligations last beyond the completion of the transaction. The acquirer's advisers, particularly law firms, may process disclosed data as independent Controllers and are then subject to the storage limitation principle in Article 5(1)(e) GDPR. Personal data may be kept only for as long as the purpose of the processing requires.[21] Frequently, a law

19 See EDPB, *Guidelines 07/2020* (footnote n 8), on the functional classification of professional service providers and processing-specific determination of roles. Classification remains fact-sensitive and may differ across activities performed by the same organisation.

20 See Articles 4(7)(8), 26, 28 and 32 GDPR; EDPB, *Guidelines 07/2020* (footnote n 8); and *Nacionalinis visuomenės sveikatos centras* (footnote n 9).

21 See Article 5(1)(e) GDPR on storage limitation and recital 39 GDPR on limiting retention to what is necessary and establishing time limits for erasure or periodic review.

firm managing the data room retains diligence materials after the transaction concludes to protect itself against professional liability claims, for example, an allegation that it failed to identify or report a material issue. That retention serves a purpose distinct from the conduct of the transaction and, as a standalone processing activity, requires its own lawful basis under Article 6 GDPR. If the files contain special categories of personal data, Article 9(2)(f) GDPR may permit their processing where necessary for the establishment, exercise or defence of legal claims.[22] Nevertheless, the appropriate retention period is consequently governed by the law firm's purpose and by any professional, statutory or regulatory obligations that apply to it. It is not determined solely by the target's retention schedule, nor may the law firm retain the full diligence record indefinitely. The retention period must remain limited to what is necessary, taking account of applicable limitation periods and professional obligations. The law firm should also consider whether parts of the record can be deleted or minimised and whether continued retention should be reviewed periodically.

Cross-border diligence presents a distinct issue. Chapter V of the GDPR applies where a Controller or Processor subject to the Regulation discloses or otherwise makes personal data available to another Controller, joint Controller, or Processor in a third country. In this context, the processing activity envisaged here is not to be confused with remote access by the same Controller who happens to be located in a third country. Such access does not, for that reason alone, constitute a transfer under Chapter V, although the associated risks remain relevant to the Controller's other obligations under the GDPR. EDPB's Guidelines 05/2021 distinguish this situation from the situations where access to personal data is provided to a third party situated in a third country, which, in such case, entails an international transfer of personal data. Where Chapter V applies, the parties must identify and rely on a valid transfer mechanism. If they rely on safeguards under Article 46 GDPR, particularly the Standard Contractual Clauses, Schrems II and subsequent EDPB guidance may require them to examine the destination country's legal environment and adopt supplementary measures where necessary.[23]

---

22 See Articles 6(1), 9(2)(f) GDPR. Article 9(2)(f) GDPR permits necessary processing of special-category data for the establishment, exercise or defence of legal claims. An independent lawful basis under Article 6(1) GDPR is still required. Applicable professional, regulatory and statutory file-retention requirements may also affect the appropriate retention period.

23 See European Data Protection Board, *Guidelines 05/2021 on the Interplay between the Application of Article 3 and the Provisions on International Transfers as per Chapter V of the GDPR* ("EDPB, *Guidelines 05/2021"*),

Transparency is another domain that presents particular concerns, as it is more difficult to reconcile with the confidentiality of the transaction that is imposed on the parties. Articles 13 and 14 GDPR continue to apply during confidential diligence, as there is no implied exemption for mergers and acquisitions, without prejudice to the situations referred to in Articles 13(4) and 14(5) GDPR. However, complying with the required transparency obligations during a confidential due diligence process may, depending on the circumstances, undermine the transaction or be difficult to implement. The solution is not to infer an unwritten M&A exemption, but to assess the transparency obligations applicable to each Controller and the specific processing involved. An existing data protection notice may satisfy those obligations if it explains, with sufficient specificity, that personal data may be processed in connection with a corporate transaction or restructuring, and if the processing in question falls within the scope of that explanation. Where the acquirer obtains personal data indirectly and becomes a Controller, Article 14 GDPR applies unless one of the exceptions in Article 14(5) is available, but those exceptions may be relied upon only when their statutory conditions are met. Data minimisation can reduce the volume of personal data subject to the transparency requirements, but it cannot remove the requirements altogether.

These considerations are not static, and the transaction structure materially influences them. In a share acquisition, the target ordinarily remains the same legal entity and Controller. The change in ownership does not itself automatically transmit personal data to the acquirer. That does not, however, eliminate transaction risk. Pre-closing disclosure of personal data to a bidder, its advisers or other participants is still a processing activity that must be assessed on its own terms. By contrast, an asset transaction may raise further disclosure and continuity questions at closing because personal data associated with transferred assets or relationships may become available to another entity, generally the acquirer.

Transaction structure therefore changes the mechanics and legal relationships through which transaction risk arises. It does not alter the need to assess each disclosure by reference to the

---

Version 2.0, adopted 14 February 2023; Case C-311/18, *Data Protection Commissioner v Facebook Ireland Ltd and Maximillian Schrems* (*Schrems II*), judgment of 16 July 2020; European Data Protection Board, *Recommendations 01/2020 on Measures that Supplement Transfer Tools to Ensure Compliance with the EU Level of Protection of Personal Data*, Version 2.0, adopted 18 June 2021; Commission Implementing Decision (EU) 2021/914 of 4 June 2021 on standard contractual clauses for the transfer of personal data to third countries [2021] OJ L 199/31.

processing actually contemplated.[24] The same governing principles apply to both transaction forms. A data room should contain no more personal data than what is necessary to answer the relevant diligence questions and assess the risks which are to be materialised as findings. That may be considerably less than the information the parties would otherwise disclose as part of a comprehensive commercial review.

IV. THE DILIGENCE INQUIRY: FROM DUE DILIGENCE TO DATA PROTECTION INTEGRATION READINESS

Conventional data protection due diligence often begins with the review of compliance artefacts, including data protection notices, records of processing activities, data protection impact assessments (DPIAs), Processor agreements, retention policies, data breach response processes, and the existence of adequate mechanisms for international transfers of personal data. These materials are necessary sources of evidence, but their existence does not establish that the processing they describe is compliant in practice. Article 5(2) GDPR makes the Controller responsible for compliance with the principles set out in Article 5(1) and requires it to demonstrate such compliance. A record maintained under Article 30 GDPR, for example, demonstrates how an organisation has documented its processing activities. It does not show whether day-to-day practice corresponds to that record. Substantive diligence should therefore examine how personal data are processed in practice and uncover differences between documented governance and operational reality.

An operational review may uncover inherited risks that remain largely invisible on paper. A lawful basis may be absent or incorrectly identified. Actual retention periods may bear little relation to the policy and fail to comply with Article 5(1)(e) GDPR. Consent methods may be deficient where consent is relied upon. Complaints, regulatory correspondence, investigations or litigation brought by data subjects may remain unresolved. Other concerns may arise from the handling of personal data breaches under Articles 33 and 34 GDPR, unaddressed security vulnerabilities, compliance with the international transfers requirements under Chapter V GDPR, or reliance on Processors, sub-Processors and undocumented legacy systems. As explained in Section 2, historical compliance findings are

24 See DSK, *Asset-Deals* (footnote n 3); see also Sections 2 and 5.

evidence of these exposures. They are not the exposures themselves. An unremediated vulnerability may present a material inherited risk even if it has not yet resulted in enforcement action.

To a certain extent, the Marriott case provides a useful and instructive illustration of these circumstances. The compromise originated in Starwood's systems before Marriott acquired the company and continued after the acquisition. The Information Commissioner expressly declined to decide whether Marriott could or should have carried out more extensive due diligence during the takeover. Instead, the enforcement action addressed Marriott's compliance after the GDPR became applicable, including whether the Starwood systems were adequately secured while those systems continued to process personal data.[25] The case does not establish a regulatory duty to undertake any particular form of mergers and acquisitions due diligence. Its relevance is narrower, and it shows that a vulnerability that originates before an acquisition does not lose its legal significance at closing if the affected processing environment remains in operation.

Even a comprehensive assessment of inherited risk remains retrospective and target-focused. It asks whether the target's current processing is lawful and adequately governed, and whether unresolved liabilities or risks remain. A transaction may, however, depend on a different question: whether that processing can lawfully become part of the acquirer's intended operating model. Conventional diligence does not necessarily provide an answer. This article uses the concept of Data Protection Integration Readiness to address that question.[26]

Data Protection Integration Readiness has two defining characteristics. First, it is prospective because it examines contemplated changes in processing, not only past or present compliance. Second, it is also relational. It does not describe an inherent quality that a target possesses in the abstract. It depends on the relationship between the target's processing environment, the particular acquirer and the future state that the acquirer intends to create. A target may be able to continue operating lawfully as a standalone business, yet cause substantial legal obstacles to an acquirer planning to centralise infrastructure, combine

25 See Information Commissioner's Office, *Marriott International, Inc., Monetary Penalty Notice* (footnote n 1), especially paras 3.2-3.6.

26 See Sherer, Hoffman and Ortiz (footnote n 2). The present article develops Data Protection Integration Readiness as a specifically legal and acquirer-relative methodology rather than as a general information-governance or systems-integration assessment.

customer datasets, introduce group-wide access or deploy common analytics. The target's present compliance position may be identical in both cases, while the legal feasibility of the contemplated future processing differs substantially.

The distinction is simple. Data protection maturity describes the target's present condition. Data Protection Integration Readiness describes whether that condition can be reconciled with the acquirer's intended future processing. A mature target may still present serious integration constraints for a particular acquirer. Conversely, a target with known deficiencies could be capable of integration once specific remedial measures have been completed. Data Protection Integration Readiness is therefore *acquirer-specific*. It cannot be reduced to a universal score assigned to the target.[27]

Data Protection Integration Readiness does not create an additional obligation under the GDPR. By contrast, this methodology applies existing obligations to the future processing environment contemplated by the transaction. Among others, purpose limitation, lawful basis, accountability, data protection by design and by default, the allocation of Controller and Processor roles, technical and organisational measures, DPIAs, and international transfer requirements already constrain what may be done with personal data after an acquisition. The value of the methodology resides in its timing. It tests those constraints whilst the intended architecture can still be changed, before integration decisions become difficult or costly to reverse.

Article 24 requires Controllers to adopt appropriate measures to ensure and demonstrate compliance. Article 25 requires appropriate technical and organisational measures both when the means of processing are determined and when processing takes place, thereby implementing the data protection by design and by default principles. The EDPB's Guidelines 4/2019 treat data protection by design and by default as continuing obligations that apply to both existing and new systems.[28] Data Protection Integration Readiness brings these

27 For complementary perspectives on integration readiness and information-systems integration in the M&A context, see Peter J Wynne, "Validating Acquisition IS Integration Readiness with Drills" (2017) *AMCIS 2017 Proceedings*, Paper 15; David Aboagye-Darko, Samuel Nii Boi Attuquayefio, Nathaniel Ankomah, Amanda Quist Okronipa and Jones Yeboah Nyame, "Information Systems Research on Mergers and Acquisitions: A Systematic Literature Review" (2024) 53(12) *Kybernetes* 5560-5581.

28 See Regulation (EU) 2016/679, arts 24 and 25; European Data Protection Board, *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default*, Version 2.0, adopted 20 October 2020. The Guidelines treat data protection by design and by default as continuing obligations and address their application to existing as well as new processing systems.

obligations into the diligence inquiry by requiring the acquirer to examine whether its proposed processing environment can satisfy those requirements before commercial assumptions harden into technical and organisational commitments.

Operationally, Data Protection Integration Readiness can be analysed through seven dimensions.

The first is processing visibility. The acquirer needs an accurate account of the personal data involved, the purposes for which they are processed, the systems in which they are held and the recipients that receive them. It must also understand the relevant Processors, international transfers, retention arrangements and material data flows. This cannot be confined to the target's formal records. An integration plan based on an incomplete picture of the target's actual processing may carry existing deficiencies into the new environment or create additional ones. Without sufficient visibility, the remaining dimensions cannot be assessed reliably.

The second dimension concerns legal continuity and purpose compatibility. The first question is whether existing processing may lawfully continue after the proposed transition. The next is whether the acquirer's intended uses are compatible with the purposes for which the personal data were originally collected. Where personal data are to be processed further for another purpose, and that processing is not based on consent or on a measure laid down by Union or Member State law, Article 6(4) GDPR requires examination of several factors. These include the relationship between the original and proposed purposes, the context in which the data were collected, the nature of the data, the possible consequences for data subjects and the safeguards available. A finding of compatibility does not dispense with the other requirements of lawful processing. It determines whether the proposed further processing can remain within the legal framework regulating the original collection.

The CJEU’s judgment in Digi shows why the new processing operation must be examined separately. The case concerned a copy of an existing subscriber database created for testing and error correction. The Court held that purpose limitation did not preclude the further processing where the new purpose was compatible with the purposes of the original collection, assessed in light of Article 6(4) GDPR. It dealt separately with storage limitation

and confirmed that the copied data remained subject to it.[29] For M&As, the main point is a limited but important one. Lawful possession of an existing dataset does not settle whether that dataset may be copied, combined or used in a different processing environment.

The third dimension is Controller and governance compatibility. Integration can change who determines the purposes and essential means of processing even if the target remains the same legal entity. A target that previously made its own processing decisions may move on to group systems, adopt purposes determined centrally, share datasets with other group entities, or participate in common processing operations. The assessment must identify the Controller or Controllers for each material processing operation or set of operations. It must then determine whether the proposed arrangement involves sole Controllership, joint Controllership under Article 26 GDPR, a Processor relationship under Article 28 GDPR or another allocation that requires clarification. The answer depends on the parties' actual functions and decision-making authority. Corporate ownership and contractual descriptions are not, in themselves, decisive for this assessment.[30]

The fourth dimension concerns the data lifecycle and the exercise of data subject rights. The target and the acquirer may use different retention periods, deletion mechanisms, transparency arrangements and procedures for responding to individuals. Consolidating their systems does not mean that data retained by the target for one purpose becomes available for unrelated use throughout the group. Migration and duplication may also make it difficult to propagate erasure, restriction, or correction across all copies held in the new environment.

The readiness inquiry should therefore determine whether retention rules, deletion mechanisms, notices and rights-handling processes can operate coherently before datasets are combined or migrated. The Digi case is relevant here as well. Even where further processing is compatible with the initial purpose, the resulting copy may be stored only for as long as necessary for the purpose for which that copy is processed.[31]

---

29 See Regulation (EU) 2016/679, Articles 5(1)(b), 6(4); Case C-77/21, *Digi Távközlési és Szolgáltató Kft v Nemzeti Adatvédelmi és Információszabadság Hatóság*, judgment of 20 October 2022, especially paras 44 and 45. The Court held that purpose limitation did not preclude the creation and storage of a separate database for testing and error correction where the further processing was compatible with the purposes of the initial collection, assessed in light of Article 6(4) GDPR.

30 See Regulation (EU) 2016/679, Articles 4(7), 26 and 28; see also the EDPB, *Guidelines 07/2020* (footnote n 8).

31 See Regulation (EU) 2016/679, Article 5(1)(e); Digi case (footnote n 29), particularly the Court's separate treatment of storage limitation after its compatibility analysis. The principles in Article 5 apply cumulatively.

The fifth dimension is technical, security and access compatibility. The fact that two systems can be connected does not establish that they should be connected from a data protection perspective. In other words, technical interoperability does not establish data protection compatibility. Integration may introduce migration tools, shared identifiers and new interfaces. It may also create duplicated data, privileged access or permissions extending across the group. Each change can alter the risk associated with processing that was previously confined to the target's environment. Articles 25 and 32 GDPR require those consequences to be assessed in relation to the proposed transition. It is not enough that the target's and acquirer's systems appeared satisfactory when reviewed separately.[32] Once databases have been combined or group-wide access has been enabled without taking into account this dimension, the material design decision may already have been implemented, leaving data protection review to remediate an existing architecture instead of shaping it from the outset.

The sixth dimension covers vendors and international transfers. The target's Processor, sub-Processor and transfer arrangements were established for its existing processing. Integration may change the Controller, the instructions given to Processors, hosting locations or the purposes for which external providers are used. It may introduce new recipients or allow personal data to be accessed from additional countries. Existing agreements under Article 28 GDPR and internal transfer mechanisms of personal data under Chapter V GDPR must therefore be tested against the proposed operating model. The relevant question is not whether the target currently possesses Processor agreements and transfer documentation. It is whether those arrangements remain valid and appropriate for the processing that will exist after the proposed transition.

The seventh dimension is remediation feasibility and sequencing. This dimension translates the previous findings identified during the due diligence phase into transactional decisions. Not every deficiency must be corrected before signing or completion. The legally relevant deadline depends on the processing operation that activates the requirement. Existing unlawful processing may call for immediate remediation or correction before completion. A constraint that arises only when databases are combined, systems are migrated, group access

32 See Regulation (EU) 2016/679, Articles 25 and 32; EDPB, *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default* (footnote n 28).

is enabled or a new analytics purpose is introduced may be addressed later. It must, however, be resolved before that particular transition takes place.

The resulting assessment should go beyond a conventional red, amber and green rating. For each material processing transition, the assessment should state whether it may proceed as planned, proceed subject to safeguards, remain segregated, be postponed pending remediation, require redesign or be abandoned. Where the proposed processing is likely to result in a high risk to the rights and freedoms of natural persons, Article 35 GDPR requires a DPIA before the processing begins. The acquisition itself does not trigger the requirement merely because it is an acquisition. The trigger is the nature and risk of the contemplated processing.[33]

This sequencing function connects the legal analysis to the execution of the transaction. A finding that a proposed processing transition cannot lawfully occur in the intended form may affect valuation, contractual protections, transitional arrangements, technical architecture, integration costs or the commercial assumptions supporting the acquisition. Conversely, a certain risk that can be addressed through a defined safeguard before a later transition need not automatically become a condition to completion. Data Protection Integration Readiness identifies both whether the operating model is legally feasible and the conditions that must be satisfied before it is implemented.

Several examples show why this assessment cannot be replaced by documentary compliance. A target may have a well-developed consent-management system, yet the consents obtained for its own services may not support the acquirer's proposed group marketing. This does not rule out the possibility that another lawful basis could apply. It means that the proposed processing activity must be assessed independently before the data are repurposed. In another example, a target may operate a legacy database whose architecture is unable to reliably propagate deletion or restriction requests across migrated copies, requiring redesign, compensating controls or replacement before particular forms of integration. A third example concerns a vulnerability originating before acquisition that continues to constrain the

33 See Regulation (EU) 2016/679, Article 35(1) and (4); Article 29 Data Protection Working Party, Guidelines *on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679*, WP248 rev.01, adopted 4 October 2017 and subsequently endorsed by the European Data Protection Board. The Guidelines identify criteria for determining whether processing is likely to result in a high risk. Lists adopted by national supervisory authorities under Article 35(4) should also be consulted.

processing environment after closing, as illustrated by the Marriott enforcement action, even without the supervisory authority having established a general rule about how far pre-acquisition diligence must extend.[34]

None of these fundamental concerns is new. Compliance gap analyses, information-system integration assessments and post-acquisition integration planning are already familiar practices. The contribution of Data Protection Integration Readiness is to connect them through a legal methodology directed at the future processing environment contemplated by a particular acquirer.[35] Integration risk is therefore not simply historical compliance risk projected forward. A target can be compliant in its existing configuration and nevertheless be unsuitable for a particular form of integration. Equally, an identified deficiency may be remediable before the processing transition to which it is relevant.

Data protection diligence must therefore examine more than the target's past and present condition. It must also determine whether the acquirer's intended processing can lawfully be built and, if so, under what conditions. That question is usually easier and less costly to answer while the future architecture remains a transaction assumption. Once it becomes an operational fact, many of the choices that mattered will already have been made.

V. SIGNING AND CLOSING: RISK TREATMENT, ALLOCATION, AND PROCESSING CONTINUITY

Identifying a data protection risk during due diligence does not determine how the transaction needs to address it. A historical infringement, an unresolved vulnerability, uncertainty about a lawful basis and a restriction on the acquirer's intended use of data do not necessarily call for the same response. Once the findings enter the negotiation of the transaction documents, the focus therefore shifts from risk identification to risk treatment. The parties must decide whether an exposure should be remedied before completion, allocated between them, reflected in the purchase price, accepted, or used to limit completion or a later integration step.

34 See Information Commissioner's Office, *Marriott International, Inc., Monetary Penalty Notice* (footnote n 1). The example illustrates the continuing relevance of an acquired security vulnerability, not a transaction-specific due-diligence obligation.

35 See Wynne and Aboagye-Darko et al (footnote n 27).

M&A practice already offers a developed set of mechanisms for doing so. Warranties may address the target's compliance, known personal data breaches, regulatory investigations, Processor arrangements and international transfers. Disclosure schedules qualify those warranties. Specific indemnities can allocate the financial consequences of an identified exposure, while covenants may require action before or after completion. Conditions precedent may prevent completion until a material deficiency has been addressed, and escrow or holdback arrangements can retain part of the consideration against identified risks.[36] These mechanisms can materially alter the parties' economic exposure and the sequencing of remediation. They cannot, however, alter obligations imposed directly by the GDPR.

That limit follows from the nature of the two regimes. The acquisition agreement allocates rights, obligations, and financial consequences between the acquirer and the seller or the target. Responsibility under the GDPR attaches to Controllers and Processors according to the functions they perform. Article 4(7) GDPR identifies the Controller by reference to who determines the purposes and means of processing, and Article 24 GDPR makes the Controller responsible for ensuring and demonstrating compliance. Supervisory authorities exercise their powers under Articles 58 and 83 GDPR by reference to infringements of the Regulation, irrespective of the allocation agreed in the transaction documents. An indemnity may require the seller or the target to reimburse the acquirer for a defined loss, but it cannot prevent enforcement against the entity responsible under the GDPR. Similarly, a warranty that no regulatory investigation is pending may give the acquirer a contractual remedy if it proves false. It does not affect the existence of the investigation or the supervisory authority's powers. Contractual arrangements, therefore, operate within the boundaries set by the GDPR and cannot displace the obligations it imposes.

The Court of Justice's judgment in Deutsche Wohnen illustrates the point in the context of administrative fines. The Court held that a fine may be imposed on a legal person acting as Controller without the infringement first being attributed to an identified natural person,

36 These mechanisms reflect established M&A risk allocation practice. Their selection depends on the nature and materiality of the identified exposure. In the data protection context, contractual protections may be combined with pre- or post-closing remediation and operational integration restrictions. See FTI Technology, *Data Risks and Challenges in M&A Transactions*, and Bird & Bird, *M&A Post Completion - Data Protection Integration and Gap Closing* (footnote n 6).

although the infringement must have been committed intentionally or negligently.[37] The judgment also separates responsibility for the infringement from the competition-law concept of an "undertaking". That concept does not determine whether, or under what conditions, a Controller is responsible. However, it becomes relevant when the administrative fine is calculated once responsibility has been established. Where the Controller forms part of an undertaking within the meaning of Articles 101 and 102 TFEU, the turnover of that undertaking is relevant to the maximum fine under Article 83(4) to (6) GDPR. In ILVA, the Court further confirmed that the economic capacity represented by the wider undertaking may also be relevant when determining an effective, proportionate and dissuasive fine.[38]

Neither judgment creates a rule of successor liability for M&A transactions. Historical GDPR liability does not pass to an acquirer merely because an acquisition has taken place. The legal effects of a transaction on pre-existing liability may depend on the target's continued legal existence, the transaction structure, applicable rules of corporate succession, and the identity of the Controller responsible for the relevant processing. In a share acquisition, for example, the target may remain in existence and continue to be the Controller responsible for an earlier infringement, even though it now forms part of a much larger undertaking. Membership of that undertaking may increase the economic weight of enforcement without turning the parent company into the infringing Controller. Responsibility and the calculation of the fine remain separate issues, as Section 2 explains in its discussion of inherited risk.

This distinction is relevant when the parties decide what to do with a due diligence finding. For transaction purposes, this distinction has practical consequences. A material historical exposure may justify further investigation, remediation, an indemnity or a valuation adjustment. In exceptional cases, it may support a condition to completion. The appropriate remediation deadline depends on the risk, rather than on the transaction timetable alone. Existing unlawful processing may require correction before completion or immediately afterwards. By contrast, an integration constraint may require no change to the target's existing operations and can instead be addressed before the specific processing transition that

37 See *Deutsche Wohnen* (footnote n 12), paras 46-52 and 61-76. The Court held that the GDPR does not require that an infringement be attributed first to an identified natural person before a legal person acting as Controller may be fined, but that an administrative fine requires intentional or negligent conduct.

38 See *Deutsche Wohnen* (footnote n 12), paras 53-59; *ILVA* (footnote n 13), paras 23-36; European Data Protection Board, *Guidelines 04/2022 on the Calculation of Administrative Fines under the GDPR*, Version 2.1, adopted 24 May 2023. See also Section 2.

activates the risk. For example, a dataset whose present use is lawful need not be remediated before closing merely because its proposed later combination with acquirer data gives rise to a purpose limitation issue. Closing need not be postponed for that reason alone. The relevant restriction is that the combination of those datasets should not occur until the legal conditions for that processing have been established. Treating closing as an artificial deadline for every remedial measure would obscure the distinction between an existing infringement and a constraint arising from a future integration step.

Completion nevertheless requires its own data protection analysis. In an asset transaction, personal data may be disclosed to another entity at closing, after which that entity begins its own processing. The target's or the seller's disclosure and the acquirer's subsequent processing are separate operations and must be assessed accordingly. The German Data Protection Conference (Datenschutzkonferenz, DSK) adopts this approach in its 2024 guidance on asset deals. It places responsibility for the disclosure on the target or the seller and responsibility for subsequent processing on the acquirer where the acquirer acts as Controller.[39] This is different from a share acquisition. If the target remains the same Controller, a change in ownership does not, by itself, result in the transmission of personal data.

As examined in Section 3, the disclosure of personal data before completion is itself subject to separate legal constraints. The DSK's 2024 guidance adopts a restrictive position in this respect, while recognising that consent or, in defined circumstances during advanced negotiations, legitimate interests may permit particular disclosures.[40] The significance for present purposes is that the lawfulness of pre-completion access and the lawfulness of personal data processing at or after completion remain two distinct questions.

The position after an asset transfer is similarly dependent on the processing concerned. Controller status should not be assumed merely because the acquirer has obtained possession of the personal data. The DSK's treatment of former-customer records provides a particularly useful illustration. Where the customer relationship has ended and records are transmitted solely to ensure statutory retention obligations can continue to be met, the DSK requires the acquirer to process those records on the target or seller's behalf under Article 28 GDPR. They

39 See DSK, *Asset-Deals* (footnote n 3), section V.1.
40 See DSK, *Asset-Deals* (footnote n 3), section I.

must be kept separate from active-customer information and, under the DSK's model, may be used by the acquirer for its own purposes only with valid customer consent.[41] In the DSK's model, this separation is described as a "two-cabinet" solution. While this position is, in part, grounded in German law and cannot be generalised to every Member State, it stresses the broader principle that acquiring personal data in this context does not, by itself, establish the lawfulness of its subsequent use.

Where the acquirer processes transmitted personal data as a Controller, additional obligations may also apply. The DSK assigns responsibility for security under Article 32 GDPR to the entity responsible for processing. Subject to the exceptions in Article 14(5) GDPR, it also requires the acquirer to provide the information prescribed by Article 14 within the applicable period after receiving the data.[42] These obligations reinforce the separation between the mechanics of the corporate transfer and the lawfulness of subsequent processing. Completion may therefore give the acquirer technical access to information while leaving the permissible purposes of its subsequent use unresolved.

Some transactions also create an intermediate processing environment between completion and full integration. Carve-outs commonly rely on transitional service arrangements under which the target or the seller continues, for a limited period, to host systems or provide services such as payroll, customer support, information technology or security. In such cases, the parties must identify their GDPR roles for each service actually provided. Article 28 GDPR applies where one entity processes personal data on behalf of another. Article 26 GDPR applies where the parties jointly determine the purposes and essential means of the same processing. These classifications depend on the functions actually performed and the influence exercised over the processing, not simply on the designation chosen in the transitional agreement.[43]

Signing and closing must, therefore, translate diligence findings into two related outputs, namely, a risk-treatment framework and a processing-transition plan. The risk-treatment framework records how identified exposures affect valuation, contractual allocation, remediation and completion. The processing-transition plan addresses the position at closing. It identifies which processing may continue, which information must remain segregated,

41 See DSK, *Asset-Deals* (n 3), section II.3.
42 See DSK, *Asset-Deals* (n 3), section V.1.
43 See EDPB, *Guidelines 07/2020* (footnote n 8).

which governance and Processor arrangements must already be in place, and which changes must wait until further legal or technical conditions have been met.

This approach shows the connection between inherited and integration risk. Signing determines how known exposures will be treated. Closing determines which processing relationships continue, change or arise at that point in the transaction. Neither determines what the acquirer may ultimately do with the personal data associated with the acquired business. That is a question of integration risk, considered in the following section. Warranties and indemnities cannot answer it, nor can the fact that the acquisition has been completed. Transaction documents can allocate and manage data protection risk, but they cannot independently confer lawfulness on the processing that the acquirer intends to undertake.

VI. POST-CLOSING AND THE INTEGRATION PROBLEM

Closing may transfer ownership of a business or its assets, but it does not expand the lawful uses of the personal data associated with them. The commercial rationale for an acquisition often points in the opposite direction. Having acquired a business partly for the value of its customer relationships, information or technology, the acquirer may expect to consolidate customer records, enrich existing profiles, centralise analytics or migrate the target's information into group systems. However, even a favourable assessment of the target's existing processing does not warrant that these later uses are lawful, since historical compliance and post-closing integration address different legal questions.

A change in ownership does not necessarily change the Controller or transfer personal data to the acquirer. In a share acquisition, the target may remain the same Controller after closing. It is worth reminding that Controller status is determined by who decides the purposes and means of processing, regardless of the corporate organisation chart.[44] Integration may later change these circumstances, as other entities may gain access, decision-making may become shared, or the parent company may determine new purposes or essential means. Data protection analysis and the identification of its obligations and risks must follow those changes as they occur.

44 See Regulation (EU) 2016/679, Article 4(7); and EDPB, Guidelines 07/2020 (footnote n 8).

Integration is consequently better understood as a series of processing transitions than a single post-closing event. Actions such as granting access to a target database, migrating data to shared infrastructure, replacing Processors, linking customer identities across businesses, and using resulting profiles for analytics may occur at different times and involve distinct obligations. The lawfulness of one transition does not guarantee the lawfulness of the following transitions. Therefore, completion cannot operate as a general legal gateway for every future use of acquired personal data.

The issue becomes especially acute when integration changes the purpose of the processing. Article 5(1)(b) GDPR requires personal data to be collected for specified, explicit and legitimate purposes and prohibits incompatible further processing. Where Article 6(4) GDPR applies to processing for a new purpose, the compatibility assessment must take account of the relationship between the original and proposed purposes, the context of collection, the nature of the data, the possible consequences for individuals and the safeguards in place.[45] This assessment forms part of the wider inquiry under Article 6 GDPR. An acquisition strategy that depends on combining or repurposing datasets must be tested against the conditions governing the proposed processing, not simply against the data's commercial value.

The Court of Justice's judgment in Digi is relevant by analogy. The case addressed the creation of an additional database containing personal data copied from an existing source. The Court treated the recording and storage of that copy as further processing and examined whether its purpose was compatible with the purposes for which the data had originally been collected. Its analysis addressed the connection between the purposes, the context of collection, the nature of the data, the possible consequences for individuals and the safeguards adopted.[46] Although Digi did not involve a merger or acquisition, its reasoning applies readily to integration. Moving, duplicating, or reusing an existing dataset in a new processing environment does not render the prior processing context irrelevant. The acquirer still needs to be able to justify the new processing and demonstrate compliance with purpose limitation and the other conditions of lawful processing.

45 See Regulation (EU) 2016/679, Articles 5(1)(b), 6 and 6(4); and recital 50.
46 See *Digi* (footnote n 29), paras 32-38.

The combination of datasets calls for careful consideration. In Opinion 06/2014, the Article 29 Working Party analysed legitimate interest processing that involves combining information from different sources or services. The analysis examined the scale and intrusiveness of the resulting processing, individuals' reasonable expectations, the impact on them, and the available safeguards during the balancing exercise.[47] Although the Opinion predates the GDPR and was not written for M&A transactions, it nevertheless captures the difficulty presented by integration strategies of this kind. Two datasets may each be processed lawfully in isolation without their combination necessarily being lawful.

The Court's judgment in Meta Platforms v Bundeskartellamt reinforces the need for a close assessment of necessity, although it does not impose a general prohibition on combining data after an acquisition. Legitimate interests can support only processing necessary for the pursued interest, and that survives the balancing of competing rights and interests. The Court also interpreted contractual necessity narrowly. The processing must be objectively indispensable to a purpose integral to the contractual relationship.[48] These requirements must be applied to the specific processing under consideration. Commercial convenience, anticipated synergies, and technical feasibility do not, alone, establish necessity under Article 6 GDPR.

Recent supervisory guidance on database reuse draws the same line. In January 2025, the French Data Protection Authority (CNIL) stated that a Controller using a database obtained from a third party must verify how it was created or disclosed and carry out its own assessment of the intended processing.[49] The recipient remains responsible for the lawfulness of its reuse. While this guidance does not specifically address mergers and acquisitions, its rationale is directly applicable to integration, as the lawfulness of the source database and the lawfulness of the recipient's intended use are connected, but separate.

On the other hand, Recital 48 GDPR does not confer on the acquirer a broader entitlement merely because the target has joined its corporate group. It recognises that Controllers within a group may have a legitimate interest in transmitting personal data internally for

47 See WP29, *Opinion 06/2014* (n 14), particularly the discussion of data combination, reasonable expectations and balancing.

48 See *Meta Platforms* (footnote n 15).

49 See Commission nationale de l'informatique et des libertés, Réutilisation de bases de données: les vérifications nécessaires pour respecter la loi, 23 January 2025.

administrative purposes, including client and employee data. This is a limited recognition, not an intra-group exemption or a blanket authorisation to consolidate data subjects' personal data. The Recital also explicitly retains the rules for third-country transfers.[50] While an interest in centralised group administration may be relevant to an Article 6 GDPR analysis, it does not mean that, for example, acquired customer profiles can automatically be repurposed for group-wide marketing, analytics, product development, or artificial intelligence training. Common ownership changes the organisational setting, but it does not convert personal data into an unrestricted group asset.

A new purpose may also trigger a separate transparency obligation. Under Articles 13(3) and 14(4) GDPR, a Controller that intends to process personal data further for another purpose must inform individuals of that purpose and supply any relevant additional information before the further processing begins.[51] The timing is critical to secure alignment with those transparency requirements, as the crucial legal event is the transition to the new purpose, not the preceding corporate closing. A general announcement of the acquisition may therefore be insufficient where subsequent integration changes the purposes, recipients, or other material features of the processing.

The 2021 decision of the Hamburg supervisory authority concerning Vattenfall shows that transparency has an independent legal role. The authority imposed a fine of €901,388.84 after approximately 500,000 individuals were not adequately informed about a data-matching exercise. It made clear that the infringement concerned the failure to provide information and did not decide whether the underlying comparison of data was lawful.[52] Although this case did not involve an acquisition, it demonstrates why transparency cannot be treated as a secondary matter once a lawful basis has been identified.

Retention presents another distinction between availability and permissible use. Information may be retained lawfully without being available for active commercial purposes. As discussed in Section 5, the German Data Protection Conference applies that distinction to former-customer records in asset transactions. Its guidance requires records transferred solely

50 See Regulation (EU) 2016/679, recital 48.

51 See Regulation (EU) 2016/679, Articles 13(3) and 14(4).

52 See Hamburg Commissioner for Data Protection and Freedom of Information, Bußgeld gegen Vattenfall Europe Sales GmbH verhängt, 24 September 2021.

to satisfy retention duties to be segregated from active customer data.[53] While this reflects merely the German supervisory position, the underlying integration question is wider. The fact that a record can technically be migrated does not mean that it may lawfully be included in the acquirer's future processing.

Infrastructure consolidation adds extra processing transitions. An acquirer may replace Processors, move the target to group cloud infrastructure, introduce common identity management or give other group entities and service providers access to the data. Some of those recipients may be located in third countries. Existing arrangements under Article 28 GDPR and Chapter V must be assessed against that future configuration. Their adequacy should not be assumed solely because they were suitable for the target's previous environment.

The application of Chapter V also depends on an accurate map of the entities involved. Under the EDPB's Guidelines 05/2021, a transfer arises when a Controller or Processor discloses or otherwise makes personal data available to another Controller, joint Controller, or Processor in a third country, provided that the Guidelines' cumulative conditions are met. The Guidelines distinguish that situation from access by an employee of the same Controller who happens to be in a third country, although the Controller's other GDPR obligations continue to apply.[54] The location from which data are accessed is therefore not enough, by itself, to determine whether Chapter V applies. The future architecture must identify both the entities involved and the capacity in which each one receives or accesses the data.

Security analysis must also concentrate on the specific transition process itself. Integration can generate temporary copies, migration environments, inherited credentials, new privileged access and interfaces between systems that were previously separate. Changes to identity and access management introduce their own risks. Those questions are not answered by establishing that the acquirer and the target each maintained adequate security before the transaction. Articles 25 and 32 GDPR apply to the design and security of the new environment. The EDPB's Guidelines on Article 25 reflect the prospective character of that obligation by treating data protection by design as relevant when selecting the means of processing.[55] By contrast, once databases are combined or group-wide access is enabled, the

53 DSK, *Asset-Deals* (footnote n 3), section II.3; see also Section 5.
54 See EDPB, *Guidelines 05/2021* (footnote n 23).
55 See EDPB, *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default* (footnote n 28).

material design decision may already have been implemented, leaving data protection review to remediate an existing architecture instead of shaping it from the outset.

The same logic applies to DPIAs. Neither the acquisition nor every integration project automatically triggers Article 35 GDPR. A DPIA is required where the proposed processing is likely to result in a high risk to individuals, having regard to its nature, scope, context and purposes. It must be completed before the processing begins.[56] Integration may introduce a risk profile quite different from the one previously assessed by the target. For example, a target might operate an ordinary customer platform and a health-related service separately, while the acquirer may intend to place both behind a common identity system, migrate them to shared infrastructure, and combine interaction data with other group information for predictive data analytics. The assessment required to identify whether Article 35 GDPR is triggered must address this future configuration. A prior DPIA does not necessarily answer the question. Article 35(11) GDPR requires a review where necessary, particularly when the risk presented by the processing changes. New purposes, broader access, increased scale, dataset combinations or a different technical architecture may therefore require the Controller to revisit and update its assessment.

Collectively, these examples show why integration risk cannot be inferred from the target's general compliance position. A target may process information lawfully, maintain a sound data protection program, and present minimal inherited risk. Yet, its legal permissions, lifecycle controls, or technical architecture may not support the processing model envisioned by the acquirer. Equally, an integration constraint may affect only one specific transition and leave the target's wider compliance unaffected.

Integration should not be approached as a systems migration to which data protection analysis is added after the technical design has been settled. Legal feasibility is part of the design of the future processing environment. Some transitions may continue without substantial

56 See Regulation (EU) 2016/679, Articles 35(1), (4) and (11); Article 29 Data Protection Working Party, *Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing Is "Likely to Result in a High Risk"*, WP248 rev.01 (footnote n 33). The Guidelines identify nine criteria for assessing whether processing is likely to pose a high risk and indicate that, in most cases, processing that meets two of those criteria should be considered to require a DPIA, although the assessment remains contextual. In addition, under Article 35(4) GDPR, most national supervisory authorities have established and published lists of the kinds of processing operations for which a DPIA is required. Those national lists should be consulted when assessing whether a contemplated processing operation is subject to that requirement.

additional measures. Others may require revised transparency measures, new Processor arrangements, transfer safeguards, technical controls, or a DPIA. Certain datasets may need to remain segregated until specific conditions are met. An intended use that cannot meet the applicable requirements may need to be redesigned or abandoned.

This is the purpose of Data Protection Integration Readiness. It offers a future-oriented approach to determine which parts of the proposed environment can lawfully be created, what conditions attach to each processing transition and where the integration strategy encounters legal or functional limits. A target may therefore be compliant in its present environment but difficult to integrate into the particular model envisaged by the acquirer. Compliance and integrability overlap, but they remain distinct concepts.

The end of the corporate transaction does not necessarily mark the end of the data protection process. In many acquisitions, the most significant data protection issues emerge precisely during the transition from ownership to operational integration.

VII. DATA PROTECTION RISK ACROSS THE M&A LIFECYCLE FRAMEWORK

The preceding analysis makes clear that data protection risk in mergers and acquisitions cannot be reduced to a single compliance inquiry conducted during due diligence. Transaction risk, inherited risk and integration risk relate to different processing operations. They arise at different times and may demand different responses. This is not because the GDPR creates a special procedure for M&A transactions. Its obligations attach to processing, not to corporate milestones such as diligence, signing, closing or integration. The lifecycle framework identifies when a transaction changes the facts relevant to those obligations and requires a legal determination at each such point. Those determinations may, in turn, constrain how the transaction proceeds. This connection between legal assessment and transactional decision-making is the principal function of the framework.

This framework is organised around five decision gates. These are analytical checkpoints, not regulatory approvals or substitutes for corporate milestones. They do not necessarily follow one another in a fixed order. Information is often incomplete during due diligence, and integration plans may continue to develop before and after completion. An initial conclusion may need to be revisited as the intended operating model becomes clearer. Each

suggested gate poses a legal question whose answer defines what the parties may lawfully do next. The suggested framework prevents the mere progression of a transaction from one corporate stage to another from being treated as justification for a material change in personal data processing.

Gate 1: What personal data may lawfully be disclosed for the transaction?

The first gate sets the lawful limits of transactional disclosure. Personal data do not become necessary for due diligence merely because they are commercially relevant. The parties must define the purpose of the proposed disclosure, identify a lawful basis and establish what personal data are actually needed for that purpose. They should also ask whether anonymised, aggregated or pseudonymised information would be sufficient, or whether less detailed information could answer the same diligence question. Relying on legitimate interests requires both necessity and the balancing exercise prescribed by Article 6(1)(f) GDPR. The processing of special categories of personal data and personal data relating to criminal convictions and offences is further subject to the additional conditions in Articles 9 and 10 GDPR.[57]

This gate should result in a reasoned decision about what may be disclosed, to whom, at which stage of the transaction and under what safeguards. Staged disclosure, clean-team arrangements, restricted access and pseudonymisation can reduce the effects of disclosure, but they cannot create a lawful basis or make unnecessary information necessary. The level of detail disclosed, i.e. the granularity of the disclosure, should be determined by the transactional question that the disclosure is intended to answer.

Gate 2: What pre-existing data protection conditions remain relevant to the acquisition?

The second gate concerns inherited risk and draws mainly on conventional data protection due diligence. The inquiry conducted during the due diligence exercise should identify, among other elements, material infringements and existing regulatory proceedings, as well as deficiencies in security, data subject requests handling, retention practices,

57 See Regulation (EU) 2016/679, Articles 5(1)(c), 6, 9 and 10; DSK, *Asset-Deals* (footnote n 3), section I; see also Section 3.

Processor arrangements, and international transfers. It should also identify any other restrictions already embedded in the target's processing operations. Due diligence should not be limited to documentary compliance, but should also examine how processing occurs in practice and establish who actually determines the purposes and means of material processing activities, since Controller and Processor status follows the functions performed.[58]

The findings then need to be characterised individually. A historical infringement that has ended is not equivalent to ongoing unlawful processing. An unresolved security vulnerability raises one kind of exposure; a restriction on the future use of a dataset raises another. The concept of inherited risk does not imply that the acquirer automatically assumes the target's historical GDPR liability. The concept covers pre-existing conditions whose legal, economic or operational consequences remain relevant to the acquisition.

The transaction can nevertheless change the economic impact of those conditions. If the Controller responsible for an infringement becomes part of an undertaking for the purposes of Article 83 GDPR, the turnover of that undertaking may be relevant to the calculation of an administrative fine. That affects the amount of the fine, not the identification of the entity responsible for the infringement.[59] The second gate should therefore produce a differentiated account of the target's existing data protection position and the effect of each finding on the transaction. A generic compliance score would conceal the distinctions that matter.

Gate 3: How should each identified risk be treated, and when?

At the third gate, each identified exposure is translated into a decision for the parties on how the transaction ought to address it. The parties must decide whether the risk should be remediated, accepted, reflected in the valuation or allocated by contract. They may also address it through a covenant or, if the risk is sufficiently material, make its resolution a condition to completion.

There are, however, two constraints that influence this decision. First, contractual arrangements may allocate economic consequences between the parties, but they cannot alter the responsibilities imposed directly by the GDPR. A warranty, indemnity, or price adjustment may regulate exposure between the acquirer and the target, but it does not,

58 See Regulation (EU) 2016/679, Articles 4(7), 24, 28 and 32; EDPB, *Guidelines 07/2020* (footnote n 8).
59 See *Deutsche Wohnen* (footnote n 12); *ILVA* (n 13); see also Sections 2 and 5.

however, resolve the separate question of whether the underlying processing complies with data protection law.

Second, the timing of remediation should be tied to the processing that creates the risk, and not to the transaction timetable. Existing unlawful processing may require correction before completion or immediately afterwards. An integration constraint may leave the target's present processing untouched for a period, but it must be resolved before the parties implement the transition to which it relates.

The output from the third gate should therefore produce more than a statement that remediation is required. It should identify the action to be taken, assign a deadline for the relevant processing, and specify the conditions that must be met before that processing may continue or change.

Gate 4: What processing may lawfully continue when the transaction closes?

The fourth gate addresses continuity. It examines what completion changes from a data protection perspective and which processing operations may continue on their existing basis. The answer depends on the structure of the transaction. In a share acquisition, the target may continue to operate as the same legal entity and Controller. The transfer of its shares does not, by itself, necessarily produce a transfer of personal data. In contrast, an asset transaction may instead involve disclosure to another entity, followed by processing by the recipient, usually the acquirer. The German Data Protection Conference's 2024 guidance illustrates the importance of separating these operations. It treats the target's disclosure and the acquirer's subsequent processing as distinct and differentiates between categories of data according to the circumstances of the transfer.[60]

Transitional service arrangements add another layer to these considerations. One party may process personal data on behalf of another, bringing Article 28 GDPR into play. Other arrangements may entail joint determination of the purposes and essential means of processing and therefore fall under Article 26 GDPR. The description chosen in the transaction documents is not conclusive. Classification depends on what the parties actually do and on the influence each exercises over the processing. The European Data Protection

60 See DSK, *Asset-Deals* (footnote n 3).

Board applies the same functional approach to Controller and Processor status as analysed before.[61]

Ideally, the fourth gate should produce a processing-continuity map. It should distinguish operations and relationships that remain unchanged at completion from those that change at that point. The map should also identify any new Controller, joint-Controller, or Processor arrangements; information that must remain segregated; and processing that cannot continue on the basis originally assumed.

Gate 5: Which intended integration transitions are legally feasible?

The fifth gate asks whether the processing environment planned by the acquirer can lawfully be built. Conventional due diligence cannot answer that question by examining the target alone. It requires Data Protection Integration Readiness, as the inquiry moves beyond the compliance of the target's existing processing and examines whether each material transition required for the future operating model can lawfully take place.

The analysis must address each transition separately. For instance, combining or repurposing datasets may require an assessment of the lawful basis and, where applicable, compatibility analysis required under Article 6(4) GDPR. Further processing for a different purpose may also trigger additional transparency requirements under Articles 13(3) and 14(4) of the GDPR. A change of Processor may require revised arrangements under Article 28 GDPR. If the future architecture entails transfer of personal data to a third country, the requirements of Chapter V of the GDPR must also be satisfied. And where processing is likely to pose a high risk to individuals, a DPIA under Article 35 GDPR may also be required before implementation. These obligations may be activated at different times and by different changes in processing. Treating the acquisition as a single data protection event risks overlooking the legal questions arising from subsequent integration steps.

Recital 48 GDPR offers only limited assistance for intra-group transmission. It recognises that transferring personal data within a corporate group for internal administrative purposes may constitute a legitimate interest within Article 6(1)(f) GDPR. As previously examined, this recital does not authorise the general combination or commercial repurposing of acquired

61 See EDPB, *Guidelines 07/2020* (footnote n 8).

datasets. Common ownership therefore does not remove the need to assess each processing activity under the GDPR.

Articles 24 and 25 GDPR make the prospective nature of this gate clear. Controllers must adopt appropriate measures to ensure and demonstrate compliance. Article 25 GDPR also requires them to consider specific data protection requirements and principles when determining the means of processing and throughout the processing itself.[62] Data protection analysis must therefore inform the design of the integrated architecture while the relevant choices remain open.

The conclusion reached at this gate need not be binary, based on an approval or rejection. One transition may proceed as planned, while another may proceed only with safeguards in place. Some data may have to remain segregated. A proposed change may need to be delayed or redesigned, and an intended use that cannot satisfy the relevant conditions may have to be abandoned. Data Protection Integration Readiness provides the method for reaching those more precise conclusions.

From legal determinations to transactional decisions

One possible objection is that the five gates turn ordinary transaction management into legal doctrine. That objection would mistake their function. A framework limited to validating, for example, the mere existence of a data protection notice, a processing agreement, or a DPIA would be largely administrative. The questions posed here are different. What personal data may lawfully enter the data room? Which processing may continue after completion? Can the proposed integration lawfully take place? These are legal questions, and the answers limit what the parties may do.

The purpose of these gates is to connect those answers with their consequences for the transaction. A disclosure decision determines what information enters the data room. Inherited exposure may affect valuation, contractual protection or remediation. A continuity problem may influence the arrangements needed at completion. An integration restriction may force a change to the technical architecture, the implementation sequence, or even a commercial assumption on which the acquisition rests. In this way, data protection analysis

62 See EDPB, *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default* (footnote n 28).

shapes the transaction without subordinating legal doctrine to transaction execution and vice versa.

The gates should also operate through a documented governance process. Material diligence findings should pass into the transaction documents and, where appropriate, into a processing-transition plan. Integration teams need to know which changes have been legally cleared, which remain conditional and which require a fresh assessment before implementation. The accountability principle under Article 5(2) GDPR, together with Articles 24 and 25 GDPR, supports this approach, as it requires that decisions be demonstrable and that data protection must be demonstrable, and data protection must influence the design of processing while the relevant choices are still being made.[63]

Taken together, these five gates give operational effect to the article's central proposition. As argued throughout this article, data protection risk in an acquisition is not a one-off exercise. By contrast, in the context of acquisitions, risk changes as personal data are disclosed, existing processing is acquired or continued, and the target's environment is transformed. Disclosure should be limited to what is necessary, while pre-existing conditions should inform valuation and risk treatment. At completion, the parties should have a clear account of which processing activities will continue, which will change and who will be responsible for them. Integration should, therefore, proceed only where the proposed future processing operations are legally and operationally feasible.

In this sense, Data Protection Integration Readiness supplies the prospective link between what the acquirer learns about the target and what it plans to build. Here, it is important to consider once again that the authority to implement a processing transition does not make that transition lawful. Instead, the transition should occur only after the applicable legal and functional conditions have been satisfied.

63 See Regulation (EU) 2016/679, Articles 5(2), 24 and 25. Note that Article 5(2) establishes the accountability principle, while Article 24 requires the Controller both to ensure and to be able to demonstrate compliance. Article 25 requires appropriate technical and organisational measures both when determining the means of processing and during processing.

VIII. CONCLUSION

Data protection in M&A cannot be reduced to a single due-diligence exercise. A transaction may generate a sequence of materially different processing operations and transitions. Personal data may be disclosed during diligence, made accessible to different parties at completion, retained in existing systems, migrated to new ones, combined with other datasets or used for new purposes after closing. These are materially different processing operations. Transaction stages do not, by themselves, determine the applicable GDPR obligations. The proper unit of legal analysis is the processing operation, or set of operations, concerned. Each must be assessed by reference to the relevant Controller or Controllers, the purposes pursued, the lawful basis relied upon and any other applicable conditions. The three-part taxonomy developed in this article, transaction risk, inherited risk and integration risk, separates questions that conventional diligence may otherwise conflate, while the five-gate framework identifies the legal determinations that should be made before the corresponding processing transitions take place.

The central proposition is simple. Closing changes the parties' corporate position, but it does not by itself expand the lawful uses of personal data associated with the acquired business. Authority over a company or its assets is distinct from authority to process the personal data connected with them. A share acquisition does not automatically change the Controller merely because ownership changes, while an asset acquisition may involve disclosures and new processing by a different Controller, each of which requires its own legal assessment. The same applies to any later migration, combination or repurposing of personal data. A favourable report on the target's historical compliance cannot be treated as permission to absorb its data into the acquirer's environment.

This is also what separates data protection maturity from Data Protection Integration Readiness. Maturity concerns the target's existing processing environment. Integration Readiness concerns whether that environment can lawfully enter the particular future state contemplated by the acquirer. The answer cannot be given in the abstract. A target may comply with the GDPR in its current configuration yet be unsuitable for the proposed integration. Conversely, a known deficiency may be remediable before the transition to which it matters. The relevant question brought by Data Protection Integration Readiness is

therefore not simply whether the target is compliant, but whether, and under what conditions, the intended future processing can lawfully be built.

For transaction advisers, this changes both the timing and the function of data protection analysis. Diligence should identify existing exposure, but it should also test the legal feasibility of the operating model on which the acquisition is premised. Identified risks may affect valuation, warranties, indemnities, covenants, conditions to completion, transitional arrangements or the sequencing of integration. Contracts can distribute financial consequences between acquirer and the seller. They cannot, however, modify the obligations imposed by the GDPR or make an otherwise unlawful processing operation lawful. Equally, remediation should not be tied mechanically to closing. Existing unlawful processing may require action before completion or immediately afterwards, while a restriction created by a planned integration step must be addressed before that step takes place.

The legal principles used in this analysis are established. The distinction between share and asset transactions, the rules governing disclosure and further processing, functional Controllership, purpose limitation, accountability, security and international transfers all derive from existing law, supervisory guidance and practice. So does the need to examine data protection during post-acquisition integration. The main contribution is methodological. The article organises these principles around the processing transitions resulting from an acquisition and connects the target's current compliance position to the legal feasibility of the acquirer's intended future state. Data Protection Integration Readiness provides the prospective inquiry. The five gates then convert that inquiry into decisions that can constrain disclosure, completion and integration before the relevant processing occurs.

The terminology of the framework is ultimately less important than the distinction it captures. An acquisition may give the acquirer corporate authority over the business, access to its systems and possession of the personal data contained in them. None of those facts determines what the acquirer may lawfully do with the personal data. The decisive question is whether the processing transition the parties intend to implement satisfies the conditions imposed by data protection law. In that sense, closing marks the completion of the corporate transaction, but not the completion of the data protection analysis.

BIBLIOGRAPHY

A. LEGISLATION AND TREATIES

- Consolidated Version of the Treaty on the Functioning of the European Union [2016] OJ C 202/47.
- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L 119/1.
- Commission Implementing Decision (EU) 2021/914 of 4 June 2021 on standard contractual clauses for the transfer of personal data to third countries pursuant to Regulation (EU) 2016/679 [2021] OJ L 199/31.

B. CASE LAW

- *Data Protection Commissioner v Facebook Ireland Ltd and Maximillian Schrems* (Case C-311/18), judgment of 16 July 2020, ECLI:EU:C:2020:559.
- *Digi Távközlési és Szolgáltató Kft v Nemzeti Adatvédelmi és Információszabadság Hatóság* (Case C-77/21), judgment of 20 October 2022, ECLI:EU:C:2022:805.
- *Meta Platforms Inc and Others v Bundeskartellamt* (Case C-252/21), judgment of 4 July 2023, ECLI:EU:C:2023:537.
- *Nacionalinis visuomenės sveikatos centras prie Sveikatos apsaugos ministerijos v Valstybinė duomenų apsaugos inspekcija* (Case C-683/21), judgment of 5 December 2023, ECLI:EU:C:2023:949.
- *Deutsche Wohnen SE v Staatsanwaltschaft Berlin* (Case C-807/21), judgment of 5 December 2023, ECLI:EU:C:2023:950.
- *Koninklijke Nederlandse Lawn Tennisbond v Autoriteit Persoonsgegevens* (Case C-621/22), judgment of 4 October 2024, ECLI:EU:C:2024:858.
- *Criminal proceedings against ILVA A/S* (Case C-383/23), judgment of 13 February 2025, ECLI:EU:C:2025:84.
- *European Data Protection Supervisor v Single Resolution Board* (Case C-413/23 P), judgment of 4 September 2025, ECLI:EU:C:2025:645.

C. EUROPEAN AND NATIONAL REGULATORY AND SUPERVISORY MATERIALS

European Data Protection Board

- European Data Protection Board, *Statement on Privacy Implications of Mergers*, adopted 19 February 2020.
- European Data Protection Board, *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default*, Version 2.0, adopted 20 October 2020.
- European Data Protection Board, *Recommendations 01/2020 on Measures that Supplement Transfer Tools to Ensure Compliance with the EU Level of Protection of Personal Data*, Version 2.0, adopted 18 June 2021.
- European Data Protection Board, *Guidelines 07/2020 on the Concepts of Controller and Processor in the GDPR*, Version 2.1, adopted 7 July 2021, minor corrections adopted 20 September 2022.
- European Data Protection Board, *Guidelines 05/2021 on the Interplay between the Application of Article 3 and the Provisions on International Transfers as per Chapter V of the GDPR*, Version 2.0, adopted 14 February 2023.
- European Data Protection Board, *Guidelines 04/2022 on the Calculation of Administrative Fines under the GDPR*, Version 2.1, adopted 24 May 2023.
- European Data Protection Board, *Guidelines 1/2024 on Processing of Personal Data Based on Article 6(1)(f) GDPR*, Version 1.0, adopted 8 October 2024.

Article 29 Data Protection Working Party

- Article 29 Data Protection Working Party, *Opinion 06/2014 on the Notion of Legitimate Interests of the Data Controller under Article 7 of Directive 95/46/EC*, WP217, adopted 9 April 2014.
- Article 29 Data Protection Working Party, *Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679*, WP248 rev.01, adopted 4 April 2017, revised and adopted 4 October 2017, subsequently endorsed by the European Data Protection Board.

European Data Protection Supervisor

- European Data Protection Supervisor, *Privacy and Competitiveness in the Age of Big Data: The Interplay between Data Protection, Competition Law and Consumer Protection in the Digital Economy*, Preliminary Opinion, 26 March 2014.

German Data Protection Conference

- Konferenz der unabhängigen Datenschutzaufsichtsbehörden des Bundes und der Länder (Datenschutzkonferenz), *Übermittlungen personenbezogener Daten an die Erwerberin oder den Erwerber eines Unternehmens im Rahmen eines Asset-Deals*, Beschluss vom 11 September 2024.

Commission Nationale de l'Informatique et des Libertés

- Commission Nationale de l'Informatique et des Libertés, *Réutilisation de bases de données : les vérifications nécessaires pour respecter la loi*, 23 January 2025.

Hamburg Commissioner for Data Protection and Freedom of Information

- Der Hamburgische Beauftragte für Datenschutz und Informationsfreiheit, *Bußgeld gegen Vattenfall Europe Sales GmbH verhängt*, press release, 24 September 2021.

Information Commissioner's Office

- Information Commissioner's Office, *Marriott International, Inc., Monetary Penalty Notice*, 30 October 2020.

D. ACADEMIC LITERATURE

- Aboagye-Darko, David, Samuel Nii Boi Attuquayefio, Nathaniel Ankomah, Amanda Quist Okronipa and Jones Yeboah Nyame, "Information Systems Research on Mergers and Acquisitions: A Systematic Literature Review" (2024) 53(12) *Kybernetes* 5560-5581.
- Holles de Peyer, Ben, "EU Merger Control and Big Data" (2017) 13(4) *Journal of Competition Law & Economics* 767-790.
- Sherer, James A, Taylor M Hoffman and Eugenio E Ortiz, "Merger and Acquisition Due Diligence: A Proposed Framework to Incorporate Data Privacy, Information Security, E-Discovery, and Information Governance into Due Diligence Practices" (2015) 21 *Richmond Journal of Law & Technology* 5.

- Vandendriessche, Robin and Caroline Buts, "Data Protection Considerations in Competition Law Assessments: A Qualitative Document Analysis of EU Decision Texts" (2025), *Journal of Competition Law & Economics* 1-43.
- Wasastjerna, Maria C, "The Role of Big Data and Digital Privacy in Merger Review" (2018) 14(2-3) *European Competition Journal* 417-444.
- Woulfe, Shane, *GDPR Considerations in M&A Transactions: How Can the Risk of Non-Compliance with the GDPR Be Reduced in the M&A Process from a Buyers' Perspective?* (LL.M. thesis, Tilburg University 2020).
- Wynne, Peter J, "Validating Acquisition IS Integration Readiness with Drills" (2017) *AMCIS 2017 Proceedings*, Paper 15.

E. OTHER INSTITUTIONAL AND PRACTITIONER MATERIALS

- Bird & Bird, *M&A Post Completion - Data Protection Integration and Gap Closing* (Bird & Bird, undated, accessed 6 September 2026).
- Cheng, Sonia, "Data Risks and Challenges in M&A Transactions" (FTI Technology, undated, accessed 6 September 2026).
- Knouff, Matthew, Rani Habash and Brad Janssen, "Data Privacy and Security Issues in M&A Transactions: Part One" (International Association of Privacy Professionals, 26 April 2016).
- Knouff, Matthew, Rani Habash and Brad Janssen, "Data Privacy and Security Issues in M&A Transactions: Part Two" (International Association of Privacy Professionals, 22 August 2016).
- The Sedona Conference, "Commentary on Data Privacy and Security Issues in Mergers & Acquisitions Practice" (2019), 20, *The Sedona Conference Journal* 233.
- Weiss, Justin and Marcin Czarnecki, *Privacy in M&A Transactions Playbook* (International Association of Privacy Professionals, 2 December 2021).



CYBERLAW

BY CIJIC

EDIÇÃO N.º XIV - SETEMBRO DE 2026

X ANIVERSÁRIO