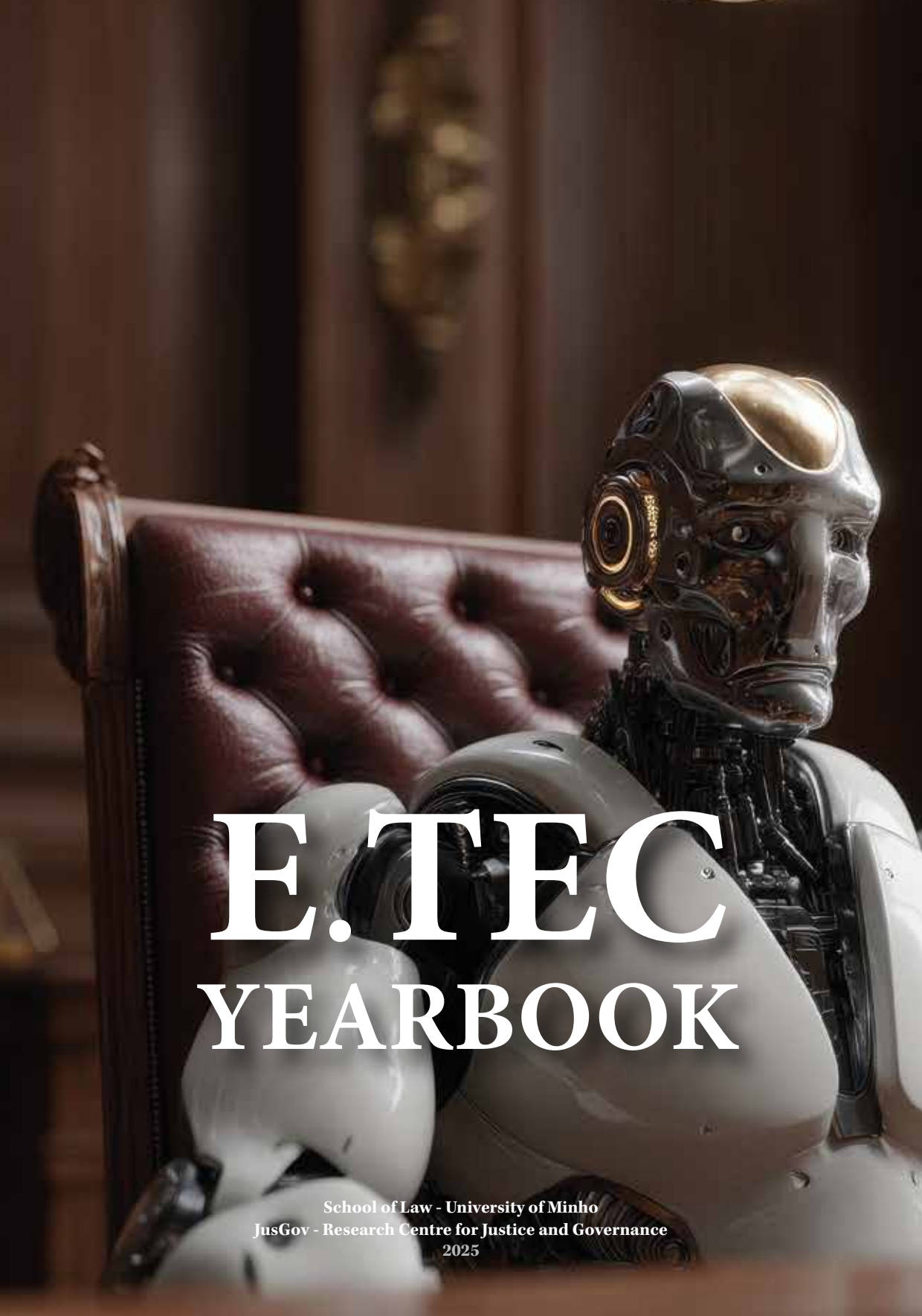

E.TEC YEARBOOK

School of Law - University of Minho
JusGov - Research Centre for Justice and Governance
2025

EDITOR
Sónia Moreira

E.Tec Yearbook

JusGov - Research Centre for Justice and Governance
School of Law - University of Minho

2025

TITLE
E.Tec Yearbook

EDITOR
Prof.ª Doutora Sónia Moreira

AUTHORS
Abner Cabral Buccini | Ana Cristina Gouveia | Ana Sirage Coimbra | Filipe Venade de Sousa Hana Horak | Luís Manuel Pica | Ng Chan Fai | Pedro Henrique Engler Urso; Matheus Raphael Pova Marinho Dantas | Sofia Ferreira da Costa

DATE
December 2025

PUBLISHERS
JusGov - Research Centre for Justice and Governance (www.jusgov.uminho.pt)
University of Minho - School of Law (www.direito.uminho.pt)

DESIGN AND TYPESET
Mariana Teixeira

COVER
Photo by Adobe Stock

ISSN
2184-707X

Financed by national funds through the National Agency for Science and Technology, FCT (Fundação para a Ciência e a Tecnologia – FCT I.P.), under the project Ref. UID/05749:Research Centre for Justice and Governance (JusGov/UM).

TABLE OF CONTENTS

PREFACE
VII

PREFÁCIO
IX

Blockchain na Administração Pública: lições desde a Estônia
Abner Cabral Buccini
1

Consentimento informado e os desafios da Inteligência Artificial
Ana Cristina Gouveia
29

Decreto-Lei n.º 2/2025, de 23 de janeiro – O regime sancionatório aplicável à violação das obrigações impostas pelo Regulamento Europeu de Governação de Dados
Ana Sirage Coimbra
51

The Framework Convention on Artificial Intelligence: An axiological compass in interpretation norms in the field of rights of people with disabilities
Filipe Venade de Sousa
75

Artificial Intelligence at the service of sustainable development and corporate governance: Brief notes on legal aspects of ESG goals and SDGs
Hana Horak
103

O princípio da legalidade, a tipicidade fiscal e a conformação de um poder discricionário da Administração Tributária em Portugal: limites funcionais à inteligência artificial
Luís Manuel Pica
121

TABLE OF CONTENTS

Da ilusão digital das deepfakes à realidade penal: a tutela da autodeterminação sexual em Macau
Ng Chan Fai
145

Neurodireitos na era digital: o conflito entre inteligência artificial e liberdade mental
Pedro Henrique Engler Urso e Matheus Raphael Pova Marinho Dantas
175

Ser idoso no século XXI: desafios e perspetivas sobre o direito à proteção da saúde na era digital
Sofia Ferreira da Costa
197

PREFACE

Welcome to the eighth E-Tec Yearbook!

The 2025 edition of the Yearbook maintains the editorial approach of previous issues, welcoming all topics related to E-Tec's core themes (Industry 4.0; Space Law and Policy; Artificial Intelligence and Robotics; New Disruptive Technological Areas for Law; and Public and Corporate Governance), giving rise to the exploration of various legal issues that emerge from technological innovations in these areas.

Thus, this volume includes very interesting and diverse texts, continuing to highlight the interdisciplinary vocation of JusGov in general and E-Tec in particular. We find texts on deepfakes and the protection of sexual self-determination; on the use of Artificial Intelligence in hiring, evaluations and dismissals; on data governance; on the protection of the elderly and the rights of people with disabilities in the digital age; informed consent and Artificial Intelligence; Blockchain in Public Administration; Neuro-rights, etc.

Once again, our sincere thanks go to the authors, the Law School of the University of Minho, in the person of its President, Professor Cristina Dias (who will soon be leaving us to take up a position in the University Rectorate), and the JusGov Board, in the person of Professor Maria Miguel Carvalho, the outgoing Director – to whom JusGov owes much of its classification in 2025 as a Centre of Excellence for Research. We would also like to take this opportunity to offer a word of encouragement to the new Director, Professor Flávia Noversa Loureiro: best wishes for your new role!

We all remain committed to our mission: to share the results of our work with the community and help legal science move forward into the future!

Braga, December of 2025.

Sónia Moreira
Coordinator/Editor

PREFÁCIO

Bem-vindos ao oitavo *Yearbook* do E-Tec!

A edição do *Yearbook* de 2025 mantém a opção editorial dos últimos números, encontrando-se aberta a todos os temas que contendem com os eixos temáticos do E-Tec (Indústria 4.0; Direito e Política Espacial; Inteligência Artificial e Robótica; Novas Áreas Tecnológicas Disruptivas para o Direito; e Governação Pública e Empresarial), dando azo à exploração de vários problemas jurídicos que emergem das inovações tecnológicas nestas áreas.

Assim, este volume integra textos muito interessantes e muito diversos, continuando a denotar-se a vocação interdisciplinar do JusGov em geral e do E-Tec em particular. Encontramos textos sobre *deepfakes* e a tutela da autodeterminação sexual; sobre o uso de Inteligência Artificial nas contratações, avaliações e despedimentos; sobre governação de dados; sobre a proteção do idoso e dos direitos das pessoas com deficiência na era digital; consentimento informado e Inteligência Artificial; Blockchain na Administração Pública; Neurodireitos, etc.

Mais uma vez, o nosso agradecimento sincero aos autores, à Escola de Direito da Universidade do Minho, na pessoa da sua Presidente, a Professora Cristina Dias (que em breve nos deixará para exercer funções na Reitoria da Universidade), e à Direção do JusGov, na pessoa da Prof.ª Doutora Maria Miguel Carvalho, a Diretora-cessante – a quem o JusGov muito deve a sua classificação em 2025 de Centro de Investigação de Excelência. Aproveitamos, ainda, para deixar uma palavra de incentivo à nova Diretora, a Prof.ª Doutora Flávia Noversa Loureiro: muitas felicidades para as suas novas funções!

Continuamos – todos – empenhados na nossa missão: partilhar o resultado do nosso labor com a comunidade e ajudar a ciência jurídica a avançar na direção do futuro!

Braga, dezembro de 2025.

Sónia Moreira
Coordenadora

BLOCKCHAIN NA ADMINISTRAÇÃO PÚBLICA: LIÇÕES DESDE A ESTÔNIA

*Abner Cabral Buccini**

Resumo: A presente pesquisa analisa o caso da Estônia, o primeiro país a utilizar a Blockchain no setor público, para compreender como essa tecnologia considerada disruptiva pode ser adotada pela Administração Pública. A Estônia é um dos países com o Estado mais tecnológico do mundo, prestando 100% dos serviços públicos de forma digital. O foco da pesquisa, portanto, é observar o fenômeno para compreender como foi a implementação da Blockchain na Administração Pública da Estônia e quais lições podem ser absorvidas. Inicia-se com o contexto histórico, partindo da Independência da Estônia no século XX. Posteriormente, conceitua-se Blockchain, apresentando suas principais características. Em seguida, explica-se os três pilares da Administração Pública digital na Estônia: e-ID, X-Road e KSI Blockchain. Para, no final, analisar os serviços públicos que, de fato, utilizam a Blockchain, os desafios legislativos e as lições que podem ser aprendidas com o exemplo da Estônia.

Palavras-chave: Blockchain; Estônia; Administração Pública.

* Mestrando em Direito Administrativo na Universidade do Minho. E-mail: pg54940@alunos.uminho.pt.

Abstract: This research analyses the case of Estonia, the first country to adopt Blockchain in the public sector, in order to understand how this disruptive technology can be applied to Public Administration. Estonia is one of the most technologically advanced states in the world, providing 100% of its public services digitally. The focus of the study is therefore to observe this phenomenon and examine how Blockchain has been implemented in Estonia's Public Administration and what lessons can be drawn from it. The research begins with the historical context, starting from Estonia's independence in the 20th century. It then defines Blockchain and presents its main characteristics. Next, it explains the three pillars of Estonia's digital public administration: the e-ID, the X-Road, and the KSI Blockchain. Finally, it analyses the public services that effectively use Blockchain, the legislative challenges, and the lessons that can be learned from the Estonian example.

Keywords: Blockchain; Estonia; Public Administration.

Introdução

Em 2008, em meio ao *crash* da bolsa de valores dos Estados Unidos e consequente crise econômica mundial, uma nova revolução financeira estava para acontecer. Seus resultados ainda demorariam para serem sentidos, porém, naquele momento, uma pessoa (ou grupo de pessoas) com o pseudônimo Satoshi Nakamoto, criavam a primeira moeda digital do mundo: o *Bitcoin*.

Atualmente, 17 anos depois, segundo a plataforma *CoinMarketCap*, existem mais de 22 mil[1] criptomoedas em circulação. O *Bitcoin* (BTC), por sua vez, continua sendo a principal e já foi adotado como reserva de valor por países como China e Estados Unidos, bem como se tornou estratégia de investimento de empresas como a *Tesla*. Em 2025, uma moeda de BTC vale aproximadamente 98 mil euros.

A criação das moedas digitais mudou a forma como o mundo compreende o dinheiro. Essas moedas permitem que pessoas transacionem sem a necessidade de um terceiro intermediário para validar as transações, pois a validação ocorre por meio do próprio sistema. Dessa forma, não há uma autoridade superior controladora da moeda. A função que por muito tempo pertenceu aos Estados e Bancos Centrais foi transmitida para o programa computacional sob o qual funciona a criptomoeda.

A revolução, no entanto, não foi somente financeira. O verdadeiro motivo do *Bitcoin* e outras criptomoedas terem permitido tal transformação no mercado econômico foi o que está por trás dele, ou seja, a tecnologia que permite ao *Bitcoin* funcionar. Como explicado, a criptomoeda garante que pessoas interessadas transacionem com maior liberdade, autonomia, transparência e economia. A razão pela qual isso é possível é a tecnologia *Blockchain*.

A *Blockchain* (cadeia de blocos), em termos simples, é um livro-razão distribuído. Assim como nas transações bancárias em que cada compra e venda de moeda é registrada em um livro-razão, nas transações de criptomoeda, esse registro é feito automaticamente em *Blockchain*. No entanto, diferentemente do livro-razão típico, a tecnologia *Blockchain* garante aos envolvidos características únicas, tais como: descentralização, pois não há um servidor central; transparência, pois todos os participantes podem visualizar as transações feitas; segurança, pois a tecnologia apresenta nível de corruptibilidade próximo a zero[2]/[3].

---

[1] EXAME, «Número de criptomoedas em circulação ultrapassa 22 mil pela primeira vez», *Exame – Future of Money*, disponível em https://exame.com/future-of-money/numero-de-criptomoedas-em-circulacao-ultrapassa-22-mil-pela-primeira-vez/ [consult. em 30/09/2025].

[2] Satoshi Nakamoto, «*Bitcoin: A Peer-to-Peer Electronic Cash System*», disponível em https://bitcoin.org/bitcoin.pdf [consult. em 30/09/2025].

[3] Menengola Everton, *Blockchain na administração pública brasileira*, Rio de Janeiro, Lumen Juris, 2022, pp. 79-80.

Essas características logo chamaram a atenção para o potencial que a tecnologia tinha, não somente para soluções financeiras, mas também para outros serviços que exigem confiança, segurança e rastreabilidade. Logo, iniciaram-se as pesquisas para sua aplicação na Administração Pública[4]. Nesse sentido, a Estônia foi pioneira, iniciando a utilização da *Blokchain* no setor público desde 2008, evoluindo até o quadro atual em que serviços básicos da sociedade, como saúde, identidade e registros públicos são realizados por meio da *Blockchain*.

A seguir, será explorado o conceito de *Blockchain* e sua adoção pela Estônia, buscando compreender, com um olhar jurídico, quais lições podem ser aprendidas da experiência estoniana. Inicialmente, será analisado brevemente a história recente da Estônia, desde as dificuldades enfrentadas no período pós-União Soviética, até se tornar um dos países com mais tecnológicos do mundo. Na sequência, será explicado como funciona, de fato, a *Blockchain* na Administração Pública na Estônia, quais os resultados e transformações que já foram obtidos e quais os desafios jurídicos e lições a aprender.

A pesquisa seguirá o método bibliográfico, com a leitura de produções acadêmicas sobre o tema, bem como nos informativos do próprio governo estoniano. Trata-se de uma pesquisa atual e fundamental que, por meio do direito comparado, busca compreender o potencial da tecnologia *Blockchain* e quais as possibilidades de aplicação em outros contextos, como o português. Nesse sentido, a própria União Europeia, por meio de estudo realizado pela Comissão Europeia, já reconheceu a importância dessa tecnologia ao afirmar que "Block-

---

[4] O termo 'Administração Pública', que por vezes pode tomar diferentes sentidos, no presente artigo é utilizado em seu sentido organizatório, com base na doutrina de José Carlos Vieira de Andrade. Como ensina Andrade, entende-se por Administração Pública em sentido organizatório "uma *estrutura complexa*, composta por um grande número de entes públicos, que se agrupam, a diversos níveis, em sectores fundamentais – na realidade há diversas 'Administrações Públicas'" (José Carlos Vieira de Andrade, *Lições de Direito Administrativo*, Coimbra, Imprensa da Universidade de Coimbra, 2017, pp. 15-17). A *Blockchain*, como será demonstrado *infra*, apresenta potencial para servir a essas diversas 'Administrações Públicas' (estaduais – direta e indireta –, autoridades reguladoras, autoridades administrativas independentes, administrações autónomas – territoriais e corporativas –, bem como entidades privadas administrativas). Por isso, o termo que se emprega no presente artigo é Administração Pública no sentido organizatório, ou seja, *lato sensu*. Dessa forma, ao analisar a utilização da *Blockchain* na Administração Pública na Estônia, o que se intenta é perceber – ainda que dentro da limitação que as poucas páginas impõem – de que modo os diversos níveis da Administração Pública da Estônia têm aplicado a tecnologia *Blockchain*. É esse também o sentido expresso pelo Governo da Estônia ao tratar sobre a organização administrativa do país no *site* do Ministério das Finanças, separando a Administração Pública em: *Central Government (State authoraties, Agencies administred by state authoraties, Public legal entities, Foundations, Central Government enterprises, Central Government non-profit organizations) e Local Government (Local government authoraties, Agencies administred by government authoraties, Foundations, Local government enterprises, Local government non-profit organizations)*, *vide*: https://www.fin.ee/en/state-local-governments-spatial-planning/riigihaldus.

chain technology can disrupt the status quo in the public sector"[5]. Portanto, é imprescindível analisar atentamente os efeitos da *Blockchain* no primeiro país que a adotou em escala nacional.

1. A implementação da *Blockchain* na Estônia

1.1. *Contexto da revolução digital*

Em 2014, a Estônia surpreendeu o mundo ao lançar o *e-Residency*, a primeira autorização de residência digital do mundo. Por meio do *e-Residency*, cidadãos de qualquer parte do mundo podem aplicar e eventualmente se tornar residentes da Estônia, usufruindo de uma série de serviços, incluindo abertura de empresa, mesmo à distância. A nova identidade digital funciona com base na tecnologia *Blockchain*, permitindo aos e-residentes autogovernarem sua identidade[6]. Além da identidade digital, atualmente a Estônia apresenta diversas soluções governamentais com base na *Blokchain*, as quais serão tratadas posteriormente e envolvem saúde, educação e até mesmo votação.

O *e-Residency* tem como objetivo atrair investidores de todo o mundo para um governo econômico e desburocratizado, mas também busca exportar as soluções em *Blockchain* para outros países interessados[7]. Por meio do *site* oficial na internet "e-Estonia", é possível acompanhar os serviços prestados através da KSI *Blokchain*. No *site*, o Governo estoniano se orgulha de ser "the first country to use Blokchain on a national level".

Quem observa a Estônia digital de 2025 não imagina a realidade vivida pelo país no século XX. Em 1940, a Estônia foi anexada pela União Soviética, ficando sob o Império Soviético até 1944, quando houve um período de ocupação da Alemanha Nazista até ao final da Segunda Guerra Mundial. Com o término da Segunda Guerra, novamente a Estônia voltou a ser parte da União Soviética. Os períodos de ocupação foram marcados por deportações em massa, mortes, pobreza e opressão[8]. Em 1990, finalmente, após uma série de eventos, foram con-

---

[5] D. Allessie; M. Sobolewski; L. Vaccari & F. Pignatelli (editor), *Blockchain for digital government*, Luxemburgo, Publications Office of the European Union, p. 11.

[6] Clare Sullivan; Eric Burger, «E-residency and Blockchain», *Computer Law & Security Review*, n.º 33, 2017, p. 475.

[7] *Ibidem*.

[8] "In the first 12 months of Soviet occupation, more than 60,000 persons were killed or deported; more than 10,000 were removed in a mass deportation during the night of June 13–14, 1941 (...) By February 1944, however, the Russians were back on the Narva front. About 30,000 Estonians escaped by sea to Sweden and 33,000 to Germany; many thousands perished at sea." In «Britanicca, Estonia – History», *Encyclopaedia Britannica*, disponível em https://www.britannica.com/place/Estonia/Independence-lost [consult. em 30/09/2025].

vocadas eleições para decidir se a Estônia se tornaria um país independente. O resultado consagrou a independência da Estônia, que foi oficialmente declarada em agosto de 1991[9]/[10].

O retorno da democracia na Estônia foi conturbado, pois o país teve de enfrentar a crise econômica e corrupção[11]. Além disso, em 2007, após divergências políticas, a tensão entre Rússia e Estônia aumentou, culminando em um *cyberattack* contra provedores de internet de bancos, empresas, Parlamento e Governo estoniano[12]. Esse fato afetou profundamente a Estônia e aumentou a preocupação com a segurança cibernética do país. A partir dali, seriam necessárias mudanças. "After receiving the cyberattack in 2007, cybersecurity became a central issue and led to implementing blockchain at an architectural level"[13].

Portanto, em cerca de 35 anos, a Estônia passou de um país em crise, recém-independente e sob constantes ameaças cibernéticas, para uma referência mundial em tecnologia para a Administração Pública. Essa transformação está totalmente relacionada com a implementação da *Blokchain* para gerar segurança, rastreabilidade e transparência. A seguir, será analisado como a Estônia estruturou parte de seus serviços estatais na *Blockchain*. Mas, primeiro, é necessário compreender qual o conceito e quais as características dessa tecnologia.

1.2. *O que é* ***Blockchain?***

Certamente, para a obtenção de uma compreensão total sobre a tecnologia *Blockchain* seriam necessárias mais páginas do que o presente artigo pode proporcionar. Dessa maneira, esse tópico não busca esgotar o tema ou fornecer informações técnicas detalhadas. O conceito será simplificado para que seja possível a operadores do direito, leigos em computação, compreenderem seu significado.

---

[9] *Ibidem*.

[10] Estonia Business and Innovation Agency, «How Estonia Regained its Independence», disponível em https://visitestonia.com/en/what-to-do/how-estonia-regained-its-independence [consult. em 30/09/2025].

[11] «Britanicca, Estonia – History», *op. cit.*

[12] Apesar de nunca ter sido confirmado, há grandes suspeitas de que o ataque tenha sido feito pela Rússia – Nato Strategic Comunications Centre of Excellence, «Hybrid Threats: 2007 cyber attacks on Estonia», disponível em https://stratcomcoe.org/publications/hybrid-threats-2007-cyber-attacks-on-estonia/86.

[13] Silvia Semenzin; David Rozas & Samer Hassan, «Blockchain-based application at a governmental level: disruption or illusion? The case of Estonia», *Policy and Society*, vol. 41(3), 2022, pp. 386-401.

Como apontado anteriormente, a *Blockchain* foi desenvolvida em 2008 por Satoshi Nakamoto para possibilitar o funcionamento do *Bitcoin*. Em seu *White Paper*, Nakamoto, explicou detalhadamente e de forma técnica o modo de execução da tecnologia, sem, contudo, mencionar uma vez sequer o nome *Blockchain*. A palavra "Block", no entanto, foi mencionada 67 vezes e a palavra "Chain" 27 vezes. O termo *Blockchain*, como utilizado atualmente, surge, então, da interpretação acerca do funcionamento da ferramenta. O emprego dessa palavra faz referência a essa cadeia sequencial e imutável de blocos que registra transações de forma distribuída e que será explicada a seguir[14].

Com base nos estudos da Comissão Europeia, define-se *Blockchain* como "um livro-razão distribuído no qual as transações de valores (em *Bitcoin* ou outro *token*) são sequencialmente agrupadas em blocos. Cada bloco está conectado ao bloco anterior e registrado de maneira imutável por meio de uma rede *peer-to-peer*[15], usando confiança criptográfica e mecanismos de garantia"[16]. A seguir, será analisado cada aspecto desse conceito para aumentar o nível de percepção sobre o assunto.

a) Livro-razão distribuído – Distributed Ledger Technology (DLT)

Inicialmente, importa entender o que é *DLT*. *Distributed Ledger Technology* (*DLT*) ou livro-razão distribuído é uma tecnologia que permite às pessoas ou computadores conectados à rede transacionarem diretamente e registrarem essas transações de forma descentralizada e sincronizada[17]. A palavra-chave para compreender essa tecnologia é: descentralização.

Enquanto, nas redes centralizadas (tradicionais), toda e qualquer transação deve passar pelo intermediário (como o Banco no exemplo das transações financeira) e registrada pelo ente centralizador; nas redes descentralizadas, não há a figura de um intermediário, pois as transações são feitas *peer-to-peer* (ponto a ponto) e os registros são realizados no sistema de maneira automática e descentralizada. Portanto, os participantes transacionam diretamente e o registro é feito automaticamente com base em criptografia. Cada ponto (computador) da

---

[14] Satoshi Nakamoto, «*Bitcoin: A Peer-to-Peer Electronic Cash System*», *op. cit.*

[15] *Peer-to-peer*, em uma tradução literal significa "ponto a ponto". Refere-se a uma rede em que cada pessoa ou dispositivo conectado respresenta ao mesmo tempo um cliente e um servidor, removendo a necessidade de um terceiro intermediário. Em suma, permite a transação direta e descentralizada.

[16] Citação retirada com tradução livre de: D. Allessie; M. Sobolewski; L. Vaccari & F. Pignatelli (editor), *Blockchain for digital government*, *op. cit.*, p. 9.

[17] *Ibidem*, p. 8.

rede tem uma "cópia" de todas as transações, que é automaticamente atualizada a cada nova transação.

Para ilustrar, observe-se a imagem a seguir que compara o método centralizado de transação/registro e o método *DLT*.

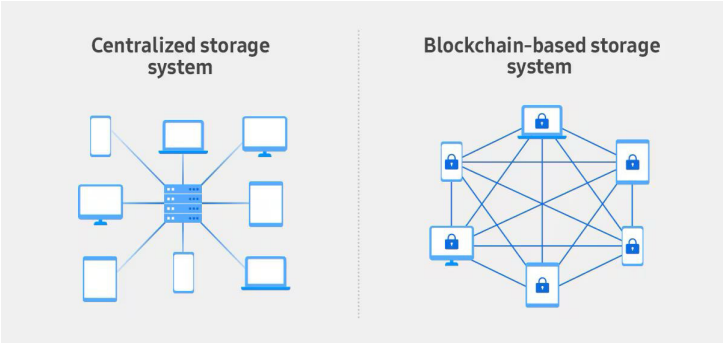

Figura 1 – Sistema centralizado X Blockchain[18]

Observa-se que no sistema centralizado as transações passam pelo intermediário para chegar ao destinatário. Já no sistema baseado em *Blockchain* (*DLT*), as transações são feitas de forma direta. Além de transações, pode-se falar em armazenamento de dados, como ilustra a imagem. No sistema centralizado, todos os dados são armazenados em um servidor central. Já no sistema *DLT* os dados são armazenados em cada computador conectado à rede, como demonstrado pela figura do cadeado. Cada aparelho conectado à rede é chamado de nó (*node*). Portanto, "ao invés de uma autoridade central manter exclusivamente a base de dados, todos os nós têm uma cópia do livro-razão, sendo que as atualizações do livro-razão da *blockchain* são propagadas através da rede em minutos ou segundos"[19].

Distributed Ledger Technology, portanto, é um gênero do qual a *Blokchain* é espécie. Assim, toda *Blockchain* é uma *DLT*, mas nem toda *DLT* é uma *Blokchain*. Existem ainda outras espécies de *DLT* com características diferentes da *Blokchchain*, mas que não serão nomeadas aqui. O que importa entender é que a *DLT* tem como principal característica registrar dados e transações de forma descentralizada.

---

[18] SAMSUNG, «Semless yet security data store», disponível em https://semiconductor.samsung.com/support/tools-resources/dictionary/blockchain/ [consult. em 30/09/2025].

[19] Relator: Ministro Aroldo Cedraz, «Levantamento da tecnologia blockchain», Tribunal de Contas da União, Brasília, TCU, Secretaria das Sessões (Seses), p. 10.

b) Transações

As transações em uma *Blockchain* podem ser de moedas, como no caso do *Bitcoin*, *Ethereum* e outras milhares de moedas digitais; de *tokens*, como no caso de NFTs ou transações imobiliárias registradas em *Blockchain* e representadas por um *token*; ou, ainda, transações de dados. Nessa última hipótese, a transação não teria como objetivo algum proveito econômico, mas sim registrar um dado na *Blokchain*, visando a segurança, a imutabilidade e a descentralização desse registro. Salienta-se que, para proteger o dado sensível, o que se registra é o *hash* desse dado, ou seja, uma sequência de números e letras que faz referência ao dado que se quer registrar.

Para a Administração Pública, o registro de dados distribuídos tem grande valor, sendo esse o maior foco dessa pesquisa. Um exemplo é a identidade digital registrada na *Blockchain*, como no caso do *e-residency*. Nesse caso, não há qualquer objetivo econômico, mas o que se busca é garantir segurança, imutabilidade e descentralização no registro dessa informação. O mesmo se aplica para o registro de um diploma universitário em *Blokchain* ou registros civis (nascimento, casamento, divórcio).

Mas porquê armazenar esses registros em um livro-razão distribuído? As principais razões são segurança e confiança. Num estudo do Tribunal de Contas da União (TCU), do Brasil, encontra-se a seguinte menção: "Enquanto um livro-razão centralizado está propenso a diversos ataques cibernéticos, um livro-razão distribuído é mais difícil de atacar, porque todas as cópias distribuídas precisam ser atacadas simultaneamente para que um ataque seja bem-sucedido"[20].

Nesse sentido, importante lembrar a intenção original do Governo da Estônia ao implementar a *Blockchain*: "the use of blockchain in Estonia is mainly driven by political priorities concerning the necessity for Building resilient infrastructures that could avoid further external intrusions"[21]. Do mesmo modo, desde a sua origem no *Bitcoin*, a questão que a *Blokchain* sempre tentou combater é o problema da confiança. Implementar a tecnologia *DLT* é retirar a confiança de terceiros intermediários e depositar no sistema computacional.

c) Cadeia de blocos

Então, como mencionado, *Blockchain* é uma cadeia de blocos que forma um livro-razão distribuído. Faz-se necessário aprofundar o entendimento sobre o que são esses blocos em cadeia. O bloco é uma estrutura que comporta dados. "Cada bloco é dividido em duas partes: cabeçalho e dados. O cabeçalho inclui

---

[20] *Ibidem*, p. 13.

[21] Silvia Semenzin; David Rozas & Samer Hassan, «Blockchain-based application at a governmental level…», *op. cit.*, p. 397.

metadados como um número único que referencia o bloco, o horário da criação do bloco e um apontador para o *hash* do bloco anterior, além do *hash* do próprio bloco"[22].

Ou seja, cada bloco tem um resumo criptográfico (*hash*[23]) do seu próprio bloco e do bloco anterior. Qualquer alteração feita em um dos blocos altera o *hash* do próximo bloco e gera uma inconsistência na rede, facilitando a identificação. Nesse sentido, fala-se em uma cadeia de blocos, pois cada bloco está interligado ao outro por meio do *hash*.

A seguir, observe-se as duas figuras do estudo do TCU brasileiro, que será de grande ajuda para compreender o funcionamento da cadeia de blocos.

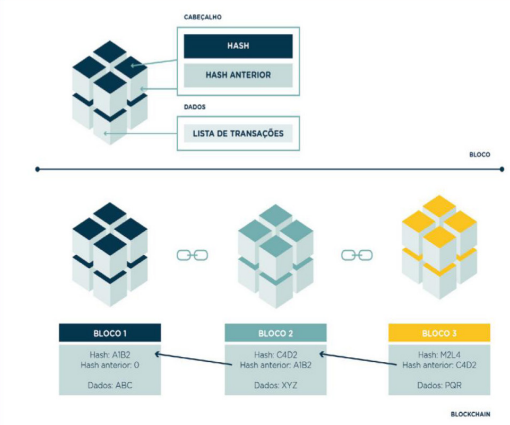

Figura 2 – Cadeia de blocos[24]

[22] Relator: Ministro Aroldo Cedraz, «Levantamento da tecnologia blockchain», *op. cit.*, p. 10.

[23] "Uma função de *hash* criptográfico, muitas vezes é conhecida simplesmente como *hash* – é um algoritmo matemático que transforma qualquer bloco de dados em uma série de caracteres de comprimento fixo. Independentemente do comprimento dos dados de entrada, o mesmo tipo de *hash* de saída será sempre um valor *hash* do mesmo comprimento." – in Kapersky, «Hash: o que são e como funcionam», disponível em https://www.kaspersky.com.br/blog/hash-o-que-sao-e-como-funcionam/2773/ [consult. em 30/09/2025].

[24] Relator: Ministro Aroldo Cedraz, «Levantamento da tecnologia blockchain», *op. cit.*, p. 10.

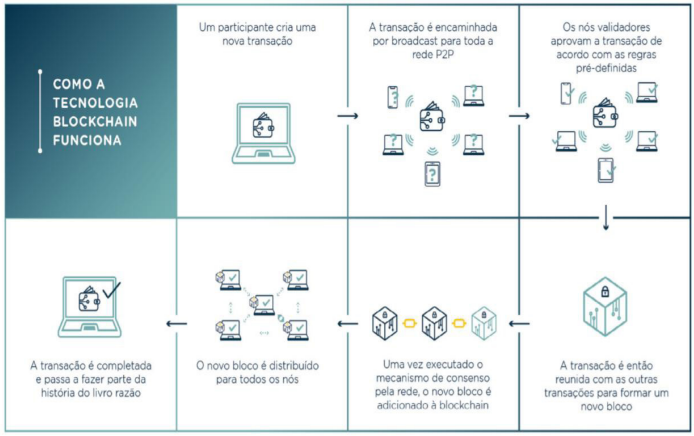

Figura 3 – Funcionamento da *Blockchain*[25]

d) Rede Peer-to-Peer

Todo esse sistema integrado forma, portanto, uma rede *peer-to-peer* (ponto a ponto) ou P2P. O termo aparece no título do *White Paper* de Satoshi Nakamoto: «A Peer-to-Peer Eletronic Cash System». "Na informática, o termo se refere a um tipo de arquitetura de rede de computadores em que cada participante (ponto) é também um servidor, e ajuda a manter o sistema funcionando"[26].

Ao contrário dos sistemas tradicionais, em que opera a lógica servidor/cliente (por exemplo no e-mail), no sistema P2P, cada nó armazena todos os dados de forma simultânea. Para ilustrar, imagine que cinco computadores estão conectados em uma rede *Peer-to-Peer* como a *Blockchain*. Se o computador 1 faz uma transação para o computador 2, todos os demais computadores terão seu sistema atualizado automaticamente e poderão conferir que a transação foi feita.

Portanto, os nós da rede são ao mesmo tempo armazenadores de dados e fornecedores de dados, no sentido de que cada transação feita por um nó alimenta a rede com novos dados. Isso poderia abrir espaço para nós mal intencionados que desejassem perverter a rede, enviando dados falsos, por exemplo. Segundo o Professor de Informática Luis C. E. Bona: "É difícil você impedir numa rede P2P o aparecimento de pessoas querendo subverter o serviço, colo-

---

[25] *Ibidem*, p. 11.

[26] «O que é Peer-to-Peer (P2P)? Entenda a relação dessa tecnologia com criptos», *Info Money*, 2022, disponível em https://www.infomoney.com.br/guias/peer-to-peer-p2p/ [consult. em 30/09/2025].

cando conteúdos que não são válidos, por exemplo"[27]. A *Blockchain* conseguiu conter parte desse problema por meio dos mecanismos de garantia que serão explicados a seguir.

e) Mecanismos de garantia (consenso)

O mecanismo de garantia ou mecanismo de consenso é o último conceito que será estudado nesse resumo sobre o funcionamento da *Blockchain*. Até o momento, já foi explicado o que é um livro-razão distribuído (*DLT*), o que são as transações, o que é a cadeia de blocos e como funciona a rede P2P. Por fim, será analisado, brevemente, o conceito de mecanismo de consenso e quais são os tipos aplicados à *Blockchain*.

Como visto, as redes P2P, por garantirem maior autonomia e descentralização para os nós, também aumentam o risco de participantes maliciosos. Para solucionar essa questão foram criados mecanismos de consenso que incentivam os participantes a trabalhar de forma honesta. O verdadeiro problema reside na validação das transações. Se um nó malicioso perverte uma transação, buscando adicionar uma informação falsa na rede, como os demais nós poderão detectar? Os mecanismos de consenso resolvem esse problema.

Portanto, os mecanismos de consenso são "as regras e os procedimentos pelos quais os nós de uma rede distribuída concordam em validar transações"[28]. Dessa forma, uma nova informação ou um novo bloco só poderá ser acrescentado à rede se passar pelo mecanismo de consenso. Em outras palavras, é o método pelo qual um novo bloco criado é aceito por todos os participantes (nós) da rede. Logicamente, todo esse mecanismo ocorre na dimensão computacional, sendo realizado pelas máquinas. Não exige necessariamente qualquer trabalho dos usuários, de modo que uma pessoa pode estar inserida em uma rede *Blockchain* sem nem saber como esses mecanismos funcionam.

Existem diversos tipos de protocolos de consenso, sendo o *PoW* (*Proof of Work*) o mais conhecido. O *PoW* é o mecanismo utilizado no *Bitcoin*, portanto, tem sua maior aplicação para *Blockchains* públicas (qualquer pessoa tem acesso). Basicamente, funciona com base na recompensa aos usuários. Quando uma nova geração de blocos é feita, cada nó (computador) ligado à rede recebe uma prova matemática computacional para resolver, com o objetivo de encontrar o valor *hash* do bloco. Nesse momento, os nós recebem o nome de mineradores, pois estão metaforicamente fazendo essa função de "minerar" até encontrar o *hash* do bloco. "O primeiro minerador que encontrar a solução, a transmite para os outros nós para que estes concordem com a solução. Assim, se o consenso for

[27] *Ibidem*.

[28] Relator: Ministro Aroldo Cedraz, «Levantamento da tecnologia blockchain», *op. cit.*, p. 14.

alcançado, o novo bloco será anexado à cadeia e ficará registrado em todos os demais blocos mantidos pela rede"[29].

Para vencer a "corrida" e solucionar a questão criptográfica, é necessário poder computacional. Portanto, esse procedimento, apesar de aumentar o nível de segurança da rede, é criticado por exigir alto consumo de energia. Um agente malicioso que tentasse perverter a rede, alterando informações, precisaria de um poder energético (e consequentemente financeiro) tão grande a ponto de conseguir refazer os cálculos para alterar cada um dos 51% dos blocos da rede, pois somente assim a rede não perceberia uma adulteração e a *Blockchain* seria atualizada conforme o novo comando malicioso. Caso contrário, os nós perceberiam a inconsistência da rede e denunciariam os blocos alterados, revertendo a alteração. Ocorre que esse é um procedimento extremamente difícil e caro de ser feito, a ponto de não compensar ao agente malicioso. Portanto, o *PoW* protege a rede, mas a um alto custo energético.

Como alternativa, foram desenvolvidos outros mecanismos de consenso que funcionam com base em regras distintas e apresentam vantagens e desvantagem em relação ao *PoW*, como o *PoS* (*Proof of Stake*), *DPoS* (*Delegated Proof of Stake*), *PoA* (*Proof os Authoraty*), entre outros.

1.3. *Principais características*

Uma vez compreendido o conceito do *Blockchain*, é possível avançar para as suas principais características, as quais fazem com que essa tecnologia seja considerada disruptiva para a Administração Pública. Como demonstrado, a *Blockchain* foi criada para resolver o problema da confiança, sendo até mesmo apelidada pela revista *The Economist* de "The Trust Machine"[30]. Portanto, a principal característica da *Blockchain* é a descentralização. Lembre-se, cada nó da rede possui uma cópia simultânea da rede, que é atualizada automaticamente, sendo possível a transação de ponto a ponto, sem a necessidade de um intermediário.

No contexto do setor público, a descentralização é importante para entregar maiores garantias aos cidadãos sobre a atuação do Poder Público. "Do ponto de vista do cidadão, o fato de poder visualizar quando quiser os dados de *blockchains* governamentais aumenta o controle social sobre as ações da Ad-

---

[29] Menengola Everton, *Blockchain na administração pública brasileira*, *op. cit.*, p. 101.

[30] «The Trust Machine», *The Economist*, 2015, disponível em https://www.economist.com/leaders/2015/10/31/the-trust-machine [consult. em 30/09/2025].

ministração Pública"[31]. Descentralizar é entregar à população mecanismos para auditar a atuação dos entes públicos.

Além da descentralização, ainda há outras características que fazem da *Blockchain* uma possível solução para a atividade da Administração Pública: a) imutabilidade – uma vez que o bloco foi aprovado por consenso, se torna imutável, pois, como demonstrado pelo protocolo de consenso, se houver alguma adulteração, esta será facilmente detectável pelos demais nós; b) segurança – os mecanismos de consenso fazem com que a rede seja segura, tornando extremamente difícil qualquer atividade maliciosa; c) transparência – como todos os nós possuem a cópia da rede, qualquer um pode conferir as transações que estão sendo feitas em tempo real; d) privacidade – nenhum dado sensível é exibido na rede, sendo substituído pelo seu *hash*, os nomes dos usuários também podem ser substituídos por pseudônimos.

Um sistema de registros que gera descentralização aumenta a confiança, traz segurança, transparência e privacidade, merece atenção. Por isso, na sequência, será analisado como funciona a *Blockchain* na Estônia, para que seja possível aprender lições com esse exemplo.

1.4. *Como a Estônia utiliza a* ***Blockchain****?*

Ainda em 2008, ano em que o *Bitcoin* foi lançado, a Estônia já havia observado o potencial da tecnologia *Blockchain* para o setor público. Naquele ano, a Estônia anunciou o projeto e-Estonia, que tinha como objetivo digitalizar todos os serviços públicos do país, utilizando a *Blockchain*[32]. Em 2024, o objetivo foi alcançado: 100% dos serviços públicos na Estonia podem ser feitos de forma *online* (inclusive casar ou se divorciar). O avanço para a digitalização já vinha sendo feito gradualmente desde o início do século XXI, porém, após os ataques de 2007, os desenvolvedores estonianos recorreram à tecnologia *Blockchain*.

O projeto e-Estonia funciona com base em três pilares: *e-ID*, *X-Road* e *KSI Blockchain*. A seguir, cada um desses pilares será explicado.

Começando pelo *e-ID*, este é, como o próprio nome sugere, uma identificação digital que todo residente na Estônia recebe, equivalente ao Cartão Cidadão português. A diferença é que o *e-ID* funciona como uma chave privada e permite ao cidadão acessar todos os serviços públicos de forma *online*.

O *X-Road*, por sua vez, é o sistema que permite às diferentes organizações públicas (saúde, educação, polícia, etc.) se interligarem e compartilharem dados

---

[31] Relator: Ministro Aroldo Cedraz, «Levantamento da tecnologia blockchain», *op. cit.*, p. 18.

[32] Silvia Semenzin; David Rozas & Samer Hassan, «Blockchain-based application at a governmental level…», *op. cit.*, p. 389.

quando necessário e autorizado. Por meio do *X-Road*, diferentes entidades estatais podem se comunicar e trocar dados. Por exemplo, quando uma criança nasce, o hospital envia os dados da criança para o departamento de Registros Públicos, que irá fazer o registro de nascimento e, em seguida, compartilhar esses dados com o departamento de Seguros, que irá liberar o benefício para a família[33].

O objetivo do *X-Road* é facilitar e desburocratizar o acesso de um departamento do Estado aos dados armazenados por outro departamento. Isso também acaba por acelerar a prestação de serviço público. Apesar de ser uma solução interessante e ter se mostrado seguro, o *X-Road* não é baseado em uma verdadeira *Blockchain*[34]. Nesse ponto reside uma das principais críticas à narrativa de utilização da tecnologia *Blockchain* na Estônia. Pois, na prática, a atuação da Administração Pública não está totalmente baseada em *Blockchain*, já que o *X-Road*, sistema que interopera os serviços, é baseado em outra tecnologia (*Markle tree*)[35]/[36]. A seguir, a figura 3 ilustra o funcionamento do *X-Road*.

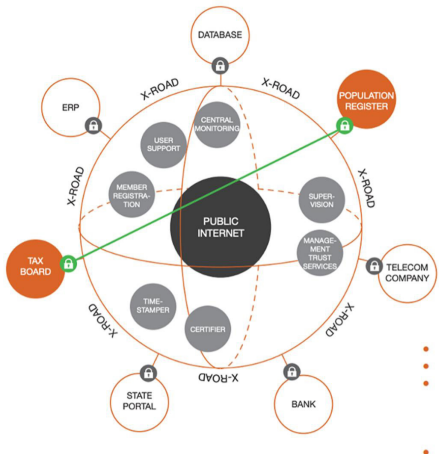

Figura 4 – *X-Road*

[33] Exemplo coletado de vídeo institucional da e-Estonia, ver em: *YouTube*, «Xroad introduction: longer version», Riigi Infosüsteemi Amet NCSC-EE, 2016, disponível em https://www.youtube.com/watch?time_continue=202&v=9PaHinkJlvA&embeds_referring_euri=https%3A%2F%2Fe-estonia.com%2F&source_ve_path=Mjg2NjY [consult. em 30/09/2025].

[34] Parol Jalakas, *Blockchain from public administration perspective*, Tallin, Tallin University of Technology, 2018, p. 24.

[35] Silvia Semenzin; David Rozas & Samer Hassan, «Blockchain-based application at a governmental level…», *op. cit.*, p. 393.

[36] Parol Jalakas, *Blockchain from public administration perspective*, *op. cit.*, p. 24.

Apesar de o *X-Road* não ser baseado em *Blockchain*, o terceiro pilar, responsável por grande parte dos serviços públicos, é o *KSI Blockchain*. Trata-se de um livro-razão distribuído, como explicado anteriormente, que foi desenvolvido na Estônia e tem como objetivo "guarantee the integrity and security of registries, identities, transactions and data privacy of its users"[37]. Diferente da *Blockchain* pública utilizada no *Bitcoin*, a *KSI* é uma *Blokchain* privada, que, por suas características, também tem um maior potencial de armazenamento de dados[38]. A Estônia utiliza a *KSI Blockchain* para registrar parte dos serviços públicos, por exemplo, prontuários de saúde, identidade digital, entre outros, o que garante maior confiabilidade e segurança.

Portanto, a digitalização do serviço público na Estônia passa por esses três pilares. Os cidadãos têm o seu próprio *e-ID* e com essa identidade acessam o serviço público de forma *online*. Por meio da *KSI Blockchain*, o Estado faz os registros de dados em *hash* referentes a determinados serviços, utilizando-se também do *X-Road* para interoperabilizar o compartilhamento de dados. Esse ecossistema digital permite que a Administração Pública seja mais segura, eficiente, econômica e célere.

Apesar de se mostrar uma solução interessante, a *Blockchain* não é a resposta para todos os problemas. Mesmo na Estônia, não são todos os serviços que utilizam a *Blockchain*. Além disso, há desafios técnicos e legislativos que precisam ser considerados. Por isso, na sequência será analisado que serviços públicos na Estônia realmente utilizam a *Blockchain* e quais os desafios legislativos para adotar essa tecnologia.

2. Adoção da *Blockchain*: serviços públicos e legislação

2.1. *Serviços Públicos via* ***Blockchain***

O grande potencial da *Blockchain* está relacionado à capacidade de registrar dados de maneira imutável, segura e descentralizada. Ocorre que o uso da *Blockchain* não pode ser feito por mero modismo, devendo ser analisado em que situações realmente há necessidade de aplicação. Ainda que a Estônia seja a maior promotora do uso da *Blockchain*, não são todos os serviços públicos do país que utilizam essa tecnologia. Apesar de 100% dos serviços serem prestados de forma *online*, o uso da *KSI Blockchain* está limitado ao sistema de saúde, com-

[37] Silvia Semenzin; David Rozas & Samer Hassan, «Blockchain-based application at a governmental level…», *op. cit.*, p. 393.

[38] ME Myrela; SNG Gourisetti & H. Culley, «Keyless Infrastructure Security Solution (KISS) – Technology Landscape Analysis Report», Pacific Northwest National Laboratory, 2018, disponível em https://www.pnnl.gov/main/publications/external/technical_reports/PNNL-27453.pdf#page=14.67 [consult. em 30/09/2025].

pra e venda de propriedades, registros de herança, sistema de Justiça digital e o departamento de mídia do Estado[39].

Em suas páginas oficiais nos *sites* da internet, a Estônia se orgulha-se de ser "o primeiro país a utilizar a *Blockchain* a nível nacional"[40]. Porém, ao aprofundar a pesquisa, é possível notar que a *Blockchain* não é utilizada em todos os serviços da Administração Pública. A pesquisa intitulada «Blockchain-based application at governmental level: disruption or illusion? The case of Estonia» inclusive critica essa apropriação do termo *Blockchain*, demonstrando se tratar principalmente de uma propaganda do Estado. Isso não anula o fato de que a Estônia foi o primeiro Estado a utilizar a *Blockchain* na Administração Pública e é uma referência no assunto. Porém, ao analisar o caso da Estônia, é necessário a ressalva de que há intenções políticas na divulgação do país como pioneiro na utilização da *Blockchain*.

Quando a Estônia se apresenta como "o primeiro país a utilizar a *Blokchain* a nível nacional" e "100% dos serviços públicos digitalizados", o que se entende é que todos os serviços utilizam a *Blokchain* como forma de registro. Porém não é o que ocorre na prática. Essa diferenciação entre propaganda e realidade serve para ensinar que, por mais que a tecnologia *Blokchain* tenha grande potencial de servir a Administração Pública, não são todos os serviços que exigem sua implementação. A *Blockchain* se apresenta como uma solução importante, principalmente, "no caso de processos que envolvam muitos agentes interessados e não confiáveis"[41]. Portanto, cabe a cada Estado, que tenha interesse em aplicar a *Blockchain*, analisar em que serviços públicos haveria maior necessidade da sua utilização.

No caso da Estônia, a principal aplicação é para os serviços de saúde. O *e-Health* foi desenvolvido em 2008 para digitalizar o sistema de saúde nacional no país. Em 2016, a empresa privada *Guardtime*, desenvolvedora da *KSI Blockchain*, fez acordo com a Estônia para a adoção da *KSI* a fim de proteger os dados da *e-Health*. A *Blockchain* permite que todos os registros de saúde feitos de forma *online* estejam protegidos e descentralizados, de modo que o paciente – dono dos seus próprios dados – possa dar permissão para médicos acessarem seus dados, além de poder sempre checar quem acessou, pois a tecnologia registra todos os acessos automaticamente.

---

[39] Silvia Semenzin; David Rozas & Samer Hassan, «Blockchain-based application at a governmental level…», *op. cit.*, p. 393.

[40] «Cybersecurity KSI Blockchain», *e-Estonia*, disponível em https://e--estonia-com.translate.goog/solutions/cyber-security/ksiblockchain/?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt&_x_tr_pto=tc [consult. em 30/09/2025].

[41] Menengola Everton, *Blockchain na administração pública brasileira*, *op. cit.*, p. 79.

Esse método de registro garante que o Estado possa escalar seu sistema de registros eletrônicos, sem precisar se preocupar com o uso indevido de dados. Ao mesmo tempo, entrega confiança ao cidadão de que somente pessoas autorizadas poderão ter acesso aos dados, sendo sempre possível conferir. Nas palavras do próprio Governo da Estônia: "The Electronic Health Record (*e-Health Record*) is a nationwide system that integrates data from Estonia's different healthcare providers to create a common record that every patient can access online. Each person in Estonia who has visited a doctor has his/her own online e-Health story that can be tracked"[42].

Para ilustrar, imagine que uma paciente natural da Estônia precisa ir ao médico por conta de uma lesão na perna. O médico local poderá acessar todos os registros pretéritos da paciente para facilitar o estudo sobre aquele caso específico. A consulta feita pelo médico local será registrada automaticamente *on-chain* (dentro da *KSI Blockchcain*). A paciente poderá conferir a qualquer momento, utilizando seu *e-ID* e acessando o *site* na internet, quem acessou seus registros e em qual horário. Todos os registros são feitos de forma descentralizada, automática, imutável e inviolável, devido à utilização da *Blockchain*. "KSI Blockchain technology is used to ensure the integrity of retrieved electronic medical records as well as system access logs"[43].

Outra aplicação da *Blokchain* conhecida na Estônia é para o *e-ID*, o qual já foi previamente conceituado na presente pesquisa. A Estônia utiliza a *Blockchain* para registrar todas as transações feitas por meio do *e-ID*. Com a identidade digital, os cidadãos da Estônia podem realizar todos os principais serviços necessários em sociedade. Utiliza-se a identidade para assinar documentos de forma digital, votar, acessar os registros de saúde, lidar com o banco e empresas, comprar e vender, entre outros[44]. Nesse contexto, a *Blokchain* é utilizada para registrar sempre que um ato é realizado com uso do *e-ID*.

Na prática, o que ocorre é o seguinte: "When a user wants to access an online service, they must first authenticate themselves using their eID card. This is done by inserting the card into a card reader and entering a PIN code. The card reader then verifies the user's digital certificate and establishes a secure connection with the online service. The Estonian government uses blockchain technology to protect the integrity of the eID system. The blockchain stores a

[42] «e-Health», *e-Estonia*, disponível em https://e-estonia.com/solutions/e-health-2/e-health-records/ [consult. em 30/09/2025].

[43] *Ibidem*.

[44] «Id Card», *e-Estonia*, disponível em https://e-estonia.com/solutions/estonian-e-identity/id-card/ [consult. em 30/09/2025].

hash of every eID transaction, which makes it impossible to tamper with or forge transactions without being detected"[45].

Dessa forma, qualquer cidadão que queira comprovar a realização de um determinado ato pode recorrer ao registro em *Blockchain*. Outras vantagens perceptíveis são: (1) segurança – pois os registros são imutáveis; (2) descentralização – o próprio cidadão pode controlar os acessos a seus registros; (3) eficiência – os serviços públicos se tornam mais céleres e menos burocráticos de serem realizados.

Portanto, o que se observa do caso da Estônia é que a implementação da *Blockchain* não se faz necessária para registrar todo e qualquer contato do cidadão com a Administração Pública, mas somente aqueles em que se percebe a necessidade de aumentar a confiança e descentralizar os registros. Nos setores em que foram implementados, a *Blockchain* parece estar sendo bem sucedida, pois se apresentou como opção escalável, atendendo mais de um milhão de pessoas[46], e segura, pois, até o momento, não apresentou qualquer incidente envolvendo o uso indevido de dados[47].

Contudo, a adoção da *Blockchain* por outros países ainda é um tema tratado com conservadorismo, pois envolve desafios que passam por desconhecimento da tecnologia, alto investimento e correspondência à legislação. Sabendo desse cenário, a União Europeia tem incentivado a produção de conhecimento e projetos nessa área. Em 2023, a Comissão Europeia lançou um *sandbox* regulatório, que terá duração até 2026 para abrir espaço para projetos que utilizam *Blokchain* no setor público[48]. No âmbito da *European Blockchain Services Infrastructure* (*EBS*), alguns projetos também têm sido realizados, envolvendo a interoperabilidade da União Europeia para reconhecimento de diplomas, registros notariais e identidade digital[49].

---

[45] «Estonia's Blockchain-Based Digital Identity System: A Model for the World», *Idefy*, 2023, disponível em https://idefy.ai/estonias-blockchain-based-digital-identity-system-a-model-for-the-world/ [consult. em 30/09/2025].

[46] Heston F. Thomas, «A case of study in blockchain health care innovation», *International Journal of Current Research*, vol 9, 2017, p. 1.

[47] Bruno Soares; António Ferreira, «The Benefits and Challenges of Blockchain Technology and eHealth Implementation in Estonia – A literature Review», *Applied Medical Informatics*, Vol. 45, No. 4/2023, p. 129 (pp. 118-131).

[48] European Commission, «Comissão lança *Sandbox* Regulatório Europeu para Blockchain», 2023, disponível em https://digital--strategy-ec-europa-eu.translate.goog/en/news/commission-launches-european-regulatory-sandbox-blockchain?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt&_x_tr_pto=tc [consult. em 30/09/2025].

[49] European Commission, «European blockchain services infrastructure», disponível em https://digital-strategy.ec.europa.eu/en/policies/european-blockchain-services-infrastructure [consult. em 30/09/2025].

A União Europeia reconhece a importância de adequar a legislação para a implementação da *Blockchain* e tem buscado compreender mais sobre a tecnologia. Na sequência, por meio da análise do caso da Estônia, serão apresentados os desafios legislativos para a implementação da *Blockchain*, principalmente no que tange à proteção de dados.

2.2. *Desafios legislativos*

O mundo dos fatos está em constante transformação, de modo que o mundo jurídico deve sempre se atualizar, para não ficar para trás, mas conseguir abordar as questões atuais, mantendo a ordem. Quando se trata da *Blockchain*, o maior desafio jurídico para sua implementação está ligado com a proteção de dados. Como discutido até aqui, a *Blockchain* tem como principal função registrar informações de forma segura, imutável, descentralizada e privada.

As atualizações legislativas que os países deverão enfrentar se quiserem implementar a *Blockchain* dependerá do ordenamento jurídico atual desse país e do setor em que se deseja inserir a tecnologia. Por exemplo, se a proposta é inserir a *Blockchain* para registros imobiliários, provavelmente leis voltadas para esse setor deverão ser atualizadas. Assim, não haverá necessidade de uma "Lei Geral *Blockchain*" para abranger todos os setores (ainda que seja interessante, dependendo do caso), se o que se deseja é a utilização para poucos e determinados setores. De qualquer modo, deverá haver atenção principalmente à proteção de dados.

Assim sucedeu na Estônia. Não houve uma lei geral para regulamentar a *Blockchain*, mas houve atualizações legislativas pontuais para os objetivos que se desejava alcançar, ao mesmo tempo que houve a observância do Regulamento Geral sobre a Proteção de Dados (RGPD) da União Europeia. Além de acolher a RGPD, a Estônia ainda publicou outras duas leis concernentes à proteção de dados, *Personal Data Protection Act (PDPA)* e *Personal Data Protection Act Implementation Act (PDPAIC)*. "The PDPA provides derogations and additional requirements to the GDPR (...). For instance, the PDPA regulates processing of personal data by law enforcement authorities for the prevention, detection, investigation and prosecution of criminal offences, as well as the imposition of punishments"[50].

Em compasso com as atualizações legislativas sobre proteção de dados, as empresas que prestam serviços de tecnologia para a Estônia devem seguir o previsto na legislação. No caso da *Guardtime*, fornecedora da *KSI Blockchain*,

[50] «Estonia», *The Global Legal Post*, disponível em https://www.globallegalpost.com/lawoverborders/data-protection-law-guide-1072382791/estonia-1067988554#1 [consult. em 30/09/2025].

por exemplo, há um *White Paper* de *compliance* explicando como a empresa se adequa às leis de proteção de dados[51]. Nesse sentido, como demonstrado pelo *White Paper*, as características da *Blockchain* parecem estar alinhadas com a maior parte dos direitos dos titulares de dados pessoais, quais sejam: direito de ser informado; direito de consentir; direito de acessar; direito de rectificação; direito de remoção; direito de portabilidade; direito de oposição[52].

Tendo em vista a natureza descentralizada da *Blockchain*, os direitos de ser informado, de acessar, de portabilidade e de oposição são amplamente assegurados, pois o titular dos dados pode acessar a qualquer momento o registro sobre quem teve acesso a seus dados, podendo monitorar por si mesmo e, se necessário, se opor a algum uso indevido. O direito de consentir também é garantido se forem aplicados mecanismos, como na Estônia, em que o titular pode autorizar e revogar a autorização de quem acessa seus dados.

Contudo, há dúvidas sobre a ampla garantia do direito de rectificação e o direito de remoção. Como visto, os registros na *Blockchain* são imutáveis e não podem ser alterados. Portanto, se algum registro for feito equivocadamente, a única forma de rectificar seria fazendo um novo registro que faz menção ao anterior, corrigindo-o indiretamente. Não seria possível alterar o registro com erro ou apagá-lo diretamente. Assim, surgem dúvidas sobre a garantia do direito de remoção, pois os registros em *Blockchain* perduram *ad eternum*.

"Quanto a este problema, há estudos sugerindo que a solução, para isso, seria salvar os dados pessoais dos usuários fora da cadeia (*off chain*)"[53]. Nesse sentido, o que se tem é um modelo híbrido, como o aplicado na Estônia, em que os dados sensíveis ficam registrados em um sistema *off chain* (fora da *Blockchain*), e o *hash* respectivo na *Blockchain* como forma de comprovar a validade desse dado. Se em algum momento for necessário apagar o dado, apaga-se *off chain* de modo a que o *hash* permaneça na *Blockchain*. Uma vez que o dado *off chain* foi apagado, se alguém tentar acessá-lo, por meio do *hash*, não terá sucesso. Dessa forma, se torna possível garantir o direito de remoção[54].

---

[51] David Shorthouse, «GDPR Compliance using KSI® Blockchain», *Guardtime, White Paper Volta*, s.d., disponível em https://m.guardtime.com/files/guardtime-whitepaper-volta-v2.pdf? [consult. em 30/09/2025].

[52] *Ibidem*.

[53] Menengola Everton, *Blockchain na administração pública brasileira*, *op. cit.*, p. 174.

[54] A mesma solução é sugerida no «Guidelines 02/2025 on processing of personal data through blockchain technologies»: "Alternatively, off-chain storage can be used, meaning that the data will, at least partially, be stored outside the blockchain, with a link or reference to the data stored in the transaction payload to address scalability or confidentiality needs." – David Shorthouse, «GDPR Compliance using KSI® Blockchain», *op. cit.*

Assim, o que se observa é que o país interessado em implementar a *Blockchain* deve fazê-lo seguindo estritamente o RGPD, bem como alterando leis específicas do setor afetado. Além disso, a União Europeia publicou o «Guidelines 02/2025 on processing of personal data through blockchain technologies»[55] com informações a respeito da adoção da *Blockchain*, de modo que qualquer Estado-membro deve seguir os preceitos estabelecidos pela União Europeia. Vale lembrar que a Estônia faz parte da União Europeia, o que demonstra ser juridicamente possível a qualquer Estado interessado implementar a *Blockchain* seguindo as diretrizes da União.

3. As 8 lições

A Estônia é o primeiro país do mundo a implementar a *Blockchain* como solução a nível nacional para serviços públicos de extrema importância, como a identidade digital e a saúde. Com a análise da experiência desse país é possível aprender algumas lições de valor para outros países da União Europeia ou fora dela que desejam implementar ou já implementam em algum nível a *Blockchain* para a Administração Pública. A seguir serão listadas oito lições principais.

3.1. *Lição 1: Não é para todos os casos*

A primeira lição revela que a *Blockchain* não é uma solução mágica para todas as esferas da Administração Pública. Nem mesmo na Estônia, país líder no assunto, a *Blockchain* é utilizada em todos os serviços públicos. Na realidade, como visto, somente alguns departamentos implementaram a *Blockchain*. As questões energética e econômica envolvendo a *Blockchain* demonstram não ser uma solução ambiental e financeiramente interessante para todos os casos. Na realidade, cabe ao Estado analisar em que setores a adoção da *Blockchain* contribuirá para a eficiência, economicidade e celeridade da Administração Pública.

Não há sentido em aplicar a *Blockchain* em setores que não a demandam. A solução se faz mais interessante para setores em que há interesses conflitantes e aumento da desconfiança ou em que o registro descentralizado é essencial. Alguns países, como o Brasil, têm estudado a implementação para coibir a corrupção em compras públicas, por exemplo. Na União Europeia, há projetos para facilitar o reconhecimento de diplomas por outros Estados-membros. Portanto, apesar do potencial disruptivo da tecnologia, deve-se compreender o contexto de cada Estado para adotá-la em setores em que haverá grande ganho.

---

[55] European Data Protection Board, «Guidelines 02/2025 on processing of personal data throughblockchain technologies», version 1.1, 2025, disponível em https://www.edpb.europa.eu/system/files/2025-04/edpb_guidelines_202502_blockchain_en.pdf? [consult. em 30/09/2025].

3.2. *Lição 2: Cooperação ao invés de substituição*

A segunda lição diz respeito à desnecessidade de excluir as tecnologias utilizadas atualmente para adotar a *Blockchain* ou demolir as estruturas administrativas existentes para implementar a *Blockchain*. Se após a análise de necessidade e adequação se optar pela adoção da *Blockchain*, deve-se estudar a possibilidade de aplicá-la dentro das estruturas já implementadas e em colaboração com as tecnologias já existentes. Aqui, não se deve adotar uma posição revolucionária ou anarquista em relação à *Blockchain*, desejando a derrubada das estruturas estatais em prol da substituição pela *Blockchain*.

O que se percebe da experiência da Estônia é que a *Blockchain* foi implementada como um *upgrade* de estruturas que já estavam em desenvolvimento. Por exemplo, a *e-Health* já existia antes da adoção da *Blockchain*. Não foi necessário criar um projeto totalmente novo, mas somente reestruturar para incluir os registros no livro-razão distribuído. Um outro exemplo é como o Brasil adotou a *Blockchain* para registros públicos sem precisar destruir os cartórios. O sistema *e-notariado* que utiliza *Blockchain* é atualmente responsável por "40% de todas as notas dos cartórios no Brasil". Portanto, a estrutura foi mantida, mas digitalizada e modernizada para gerar eficiência.

3.3. *Lição 3: Implementação gradual*

Diante de uma novidade tecnológica, sempre há a corrida para despontar na frente. Porém, a implementação da *Blockchain* no setor público deve ser gradual, como tem sido na União Europeia, sob pena de ferir o RGPD ou leis locais e gerar danos ao erário público. Hoje a Estônia pode ser considerada um dos países com o Estado mais tecnológico do mundo, com serviços públicos registrados na *Blockchain*, mas essa realidade já está sendo construída há pelo menos 25 anos. Para um país que deseja se digitalizar, como no caso de Portugal, que tem apresentado projetos no âmbito da "Estratégia para a Inovação e Modernização do Setor Público", deve existir o entendimento de que as mudanças serão graduais.

A implementação acelerada pode significar em primeiro lugar o desrespeito pelas normas de utilização da tecnologia, como a «Guidelines 02/2025 on processing of personal data through blockchain technologies», mas, principalmente, a violação do RGPD. A *Blockchain* tem como função o registro de transações, que, no âmbito da Administração Pública, envolve dados pessoais. Portanto, qualquer projeto de adoção da tecnologia deve ser feito de forma gradual.

Como no caso da Estônia, a parceria público-privada é essencial para implementar a *Blockchain*. Sem a empresa *Guardtime*, que desenvolveu a *KSI*, a Estônia não teria material para a utilização da nova tecnologia. Em outros paí-

ses, provavelmente, também será necessário comprar soluções privadas. Dessa forma, é essencial a curadoria, por meio de regras de *compliance* e observância das leis vigentes.

3.4. *Lição 4: Análise de custo-benefício*

Como visto na Lição 1, a *Blockchain* não deve ser implementada em todo e qualquer caso, mas somente naqueles em que realmente haverá proveito, ou seja, aqueles casos em que a adoção da tecnologia significará eficiência, economicidade e celeridade, nos termos da Boa Administração. Portanto, antes da adoção da tecnologia é necessário que se faça essa avaliação.

A forma como a Estônia implementou a *Blockchain* no setor público serve como aprendizado, porém não pode ser totalmente moldada e aplicada em outro contexto. A primeira preocupação é quanto à escalabilidade. Ora, a *Blockchain* no *e-Health* da Estônia alcança um pouco mais de 1 milhão de pessoas. Mas como funcionaria em Portugal, com mais 10 milhões de pessoas? Ou então, no Brasil, com mais de 200 milhões de pessoas?

Estudos demonstram que a escalabilidade da *Blockchain* é um problema a ser considerado. "(...) o modelo tradicional de cadeia única de Blockchain tende a empacotar e armazenar transações dentro de blocos na cadeia e, com o aumento de transações, o desempenho da cadeia como transação e tempo de confirmação ficam seriamente afetados"[56]. Esse problema ficou inclusive conhecido como o "trilema da *Blockchain*", conceito desenvolvido pelo criador da moeda digital *Ethereum*, segundo o qual o maior desafio da *Blockchain* consiste em alcançar ao mesmo tempo segurança, descentralização e escalabilidade[57]. Portanto, a análise minuciosa do caso concreto é necessária para evitar investimentos sem frutos.

A segunda preocupação diz respeito ao gasto energético e consequentemente econômico da implementação da *Blockchain*. Os computadores utilizados para as transações na *Blockchain* consumem muita energia, o que gera tensões ambientais e também econômicas. Isso também está conectado com o mecanismo de consenso da *Blockchain*. Nesse sentido, "(...) os sistemas híbridos PoW e PoS têm recebido bastante atenção da academia por retirarem o melhor dos dois sistemas e privilegiarem a sustentabilidade"[58].

---

[56] R. Val; L. B. Gouveia, «Escalabilidade no armazenamento na Blockchain», *Brazilian Applied Science Review*, Curitiba, vol. 7, no. 2, 2023, p. 587.

[57] «What is the blockchain trilemma», *Coinbase*, disponível em https://www.coinbase.com/pt-br/learn/crypto-glossary/what-is-the-blockchain-trilemma [consult. em 30/09/2025].

[58] Menengola Everton, *Blockchain na administração pública brasileira*, *op. cit.*, p. 111.

Dessa forma, a análise do custo-benefício na implementação da *Blockchain* deve ter em conta tanto a questão da viabilidade, que analisa a escalabilidade, quanto a sustentabilidade ambiental e econômica. Se, após a análise, se perceber que os benefícios superam os custos, então, deve-se implementar. Caso contrário, deve-se permanecer com a solução atualmente adotada ou buscar tecnologia diversa.

3.5. *Lição 5: Ordenamento jurídico*

A quinta lição retirada do caso da Estônia é que, tratando-se da União Europeia, em grande parte dos países, não há necessidade de alterar todo o ordenamento jurídico. Do mesmo modo não é necessário uma "Lei Geral *Blockchain*", bastando a atualização legislativa dos setores que receberão a nova tecnologia. Nesse sentido, é essencial que juristas, com entendimento no assunto, analisem a legislação nacional e avaliem que adaptações precisam ser feitas para receber a tecnologia *Blockchain*.

3.6. *Lição 6: Adequação ao RGPD*

A sexta lição está relacionada ao RGPD. Ainda que sejam múltiplos os desafios que envolvem Direito e tecnologia, quando se trata do estudo da *Blockchain*, há uma questão jurídica decisiva e fundamental que merece ser destacada: a proteção de dados. A digitalização da Administração Pública tem, ao mesmo tempo, o bônus de facilitar a vida dos cidadãos, mas também o ônus de utilizar os dados pessoais para que isso seja possível. Ora, um Estado, como a Estônia, em que 100% dos serviços públicos podem ser prestados de forma digital precisa do máximo de dados possível de toda a população para tornar essa realidade possível. Assim, a principal preocupação está na forma como esses dados serão utilizados.

Para tanto é essencial que os Estados e empresas prestadoras de serviços atentem ao RGPD e, inclusive, criem novas leis para aumentar as garantias dos titulares de dados frente aos abusos cometidos pela Administração. Como visto, de maneira geral, a *Blockchain* está de acordo com o RGPD, pois suas características principais garantem os direitos previstos no Regulamento. Mesmo assim, a tecnologia deve ser desenvolvida em observância a todos os aspectos legais, inclusive no que tange às questões mais complexas, como é o caso do direito à rectificação e direito de remoção.

3.7. *Lição 7:* ***Sandboxes*** *regulatórias*

Como demonstrado, a União Europeia está utilizando uma *Sandbox* regulatória para incentivar projetos envolvendo *Blockchain*. Essa é uma solução importante para implementar a *Blockchain* de forma gradual, como aprendido na Lição 3, entender na prática o custo-benefício (Lição 4) e as alterações legislativas necessárias (Lição 6). *Sandbox* faz menção às caixas de areia onde as crianças brincam soltas, sem qualquer perigo, mas com a supervisão dos pais. "Daí advém a ideia de criar espaços, onde as novas tecnologias (como as redes *blockchain*) possam ser testadas, mas em um ambiente controlado, sem riscos à Administração Pública e aos usuários"[59].

Dessa forma, seguindo o padrão estabelecido pela UE, os Estados interessados em atrair projetos que envolvam *Blockchain* podem criar suas próprias *Sanboxes* para testar a utilização da tecnologia.

3.8. *Lição 8: Investimento técnico*

A partir de toda a discussão apresentada fica clara a necessidade de investimento em conhecimento, como a União Europeia tem feito por meio dos projetos lançados recentemente. Sem conhecimento técnico, tanto informático quanto jurídico, não será possível avançar para soluções que têm potencial de transformar a Administração Pública e assegurar os direitos dos cidadãos. Para tanto, destaca-se a necessidade do intercâmbio entre operadores do direito e engenheiros da computação, para que seja possível o desenvolvimento de leis que criem o espaço adequado para o florescimento da *Blockchain*.

[59] Menengola Everton, *Blockchain na administração pública brasileira*, *op. cit.*, p. 147.

4. Considerações finais

A presente pesquisa tencionou analisar, por meio do caso da Estônia, o potencial da *Blockchain* para a Administração Pública. Com o estudo de caso, foi possível compreender que a Estônia optou por adotar a *Blockchain* para responder a um *cyberattack* em 2007. A opção pela tecnologia visou, em um primeiro momento, a segurança dos dados, já que, tratando-se de *Blockchain*, qualquer acesso fica registrado.

A jornada da Estônia em busca da digitalização fez com que, atualmente, o país alcançasse a marca de 100% dos serviços públicos digitalizados. Contudo, a utilização da *Blockchain* limita-se a algumas soluções, como *e-Health*, *e-ID* e registros públicos. Assim, estudando os aspectos técnicos da *Blockchain*, refletindo na forma de utilização da Estônia e nas regulamentações existentes, foi possível aprender lições valiosas que terão grande utilidade em futuras pesquisas ou em planejamentos para a implementação da tecnologia.

Por fim, insta salientar que a presente pesquisa não busca esgotar o tema, mas apenas contribuir para o desenvolvimento do assunto. Deve-se ressaltar que ainda há problemas em relação à *Blockchain* no setor público que precisam ser aprofundados, tais como: a questão energética; a escolha do mecanismo de consenso ideal; se a *Blockchain* deve ser pública ou privada; a atenção à proteção de dados; o direito de rectificação e direito de remoção; a questão da escalabilidade. A utilização da *Blockchain* no setor público, de maneira geral, está só começando e mais do que nunca é necessário que juristas se engajem no tema para propor soluções.

CONSENTIMENTO INFORMADO E OS DESAFIOS DA INTELIGÊNCIA ARTIFICIAL

*Ana Cristina Gouveia**

Resumo: Desde o século passado que o setor médico tem vindo a atravessar uma linha evolutiva de paulatina digitalização. Não obstante, o advento da Inteligência Artificial (IA) tornou aquela realidade manifestamente expressiva. Atualmente, a implementação do recurso e utilização de máquinas dotadas de características típicas do ser humano atravessam todo o *iter* da prestação de cuidados de saúde, desde o momento de diagnóstico patológico, à prescrição e execução do tratamento clínico. Tal panorama tem impactado de modo profundo a doutrina do consentimento informado. Este, invariavelmente, terá de se revestir de particularidades que atendam às características próprias da tecnologia em causa.

Diante disto, o presente escrito detém como propósito precípuo a análise da dinâmica do consentimento informado quando contracenado com sistemas de IA. Para tal, discorreremos sobre os benefícios e riscos da utilização da IA, a forma como o consentimento deve ser prestado e as informações a transmitir para que o mesmo se caracterize informado e esclarecido.

Palavras-chave: Consentimento informado; Inteligência Artificial; Saúde.

* Assistente Convidada na Faculdade de Direito e Ciência Política da Universidade Lusófona – Centro Universitário do Porto. Investigadora Associada do CEAD – Francisco Suárez. Doutoranda em Direito, especialidade em Direito Privado. Advogada. E-mail: p8587@ulusofona.pt.

Abstract: Since the last century, the medical sector has been undergoing a gradual process of digitization. However, the advent of Artificial Intelligence (AI) has made this reality even more evident. Currently, the implementation of this resource and the use of machines equipped with characteristics typical of human beings permeate the entire healthcare process, from the moment of pathological diagnosis to the prescription and execution of clinical treatment. This scenario has had a profound impact on the doctrine of informed consent. This will invariably have to take on particularities that meet the specific characteristics of the technology in question.

In view of this, the main purpose of this paper is to analyse the dynamics of informed consent when contrasted with AI systems. To this end, we will discuss the benefits and risks of using AI, how consent should be given, and the information that must be provided for consent to be considered informed and clear.

Keywords: Informed consent; Artificial Intelligence; Health.

1. Introdução

A IA configura, hoje, um marco assinalável no seio da quarta revolução industrial. Afastada que foi de mero cenário da ficção científica, tendo-se tornado, pelo invés, grande parte da realidade, esta emergente e disruptiva tecnologia tem transformado de modo múltiplo e transversal a sociedade. O escopo da prestação de cuidados de saúde não configurou exceção. Constituindo-se, aliás, como uma das áreas mais promissoras, existem já autores a colocar em linha de conta a altura em que a máquina inteligente irá, finalmente, ganhar terreno sobre a antiga profissão do médico, chegando a substituí-la.

Imagine-se um caso em que um paciente é diagnosticado pelo médico com uma determinada doença e, em favor da mesma, recomenda um tratamento específico. Para o efeito, explicou ao paciente as etapas do tratamento e recuperação, bem como os seus benefícios e riscos. O paciente consentiu no tratamento recomendado pelo médico sem saber, contudo, que o tratamento sugerido não resultou de uma apreciação sua, mas de uma sugestão apresentada por um sistema de IA, a partir da análise dos dados do paciente. Pois bem, na senda de GLENN COHEN, "Well, come to think of it, do you know whether an AI/ML system, *was* used last time you saw a physician?"[1].

É inegável que os significativos avanços da tecnologia têm permitido alcançar mais e melhores resultados. Todavia, numa conjuntura progressivamente mais inteligente, a prestação de cuidados médicos não pode conhecer desvios frente aos princípios basilares outrora erguidos, mormente, o da autonomia do doente, corolário do consentimento informado. Nesta seara, terá o prestador dos cuidados de saúde de informar o doente da presença, recurso e utilização de um sistema de IA? Em caso afirmativo, exigir-se-á igualmente uma explicação da forma de funcionamento geral do sistema, ou dos parâmetros em que concretamente se baseou a decisão?

Sob este novo e admirável mundo inteligente e disruptivo, o propósito do nosso estudo é confrontar as questões colocadas – para cujas respostas se encontram longe de ser simples – com as soluções oferecidas pela ciência jurídica. No cerne da contenda encontra-se, pois, como dilemática de grande monta, a de saber como configurar a autonomia do doente frente à possibilidade de ser tratado por um sistema de IA. Em ordem à mesma, é nossa intenção descortinar as formas de introdução da IA em contexto sanitário, bem como, de forma sistemática e pormenorizada, dentro dos limites que se nos impõe no presente escrito, a doutrina do consentimento informado. Tudo sem nunca perder de vista a

[1] Glenn Cohen, «Informed Consent and Medical Artificial Intelligence: What to Tell the Patient?», *The Georgetown Law Journal*, Volume (Vol.) 108, 2020, página (p.) 1426, disponível em https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3529576 [consultado em 18/07/2025].

interseção dos tópicos enunciados num contexto sanitário, dentro de um quadro legal europeu recentemente desenhado a propósito da temática.

Note-se que, no seio de uma linha evolutiva de profunda transformação da prática da medicina, esta é, somente, uma das problemáticas que a introdução da IA conforma. Contudo, é uma das vertentes sobre as quais parcamente se escreveu. Assim, enformando-se a temática de cariz moderno e iminente, com uma complexidade comedidamente descortinada pela doutrina, a nossa investigação será pioneira, resultando demarcada a sua pertinência.

2. O consentimento

Vai longe a conceção clássica do médico, dono do conhecimento, diante de um doente incapaz, enfermo, desprovido de perseverança de julgamento ou de vontade[2]. Afastado que foi o paternalismo milenar de Hipócrates que adjudicava àquele primeiro o poder de bem decidir em nome do segundo, contemporaneamente, a prestação de cuidados de saúde – seja de diagnóstico ou de tratamento – pressupõe a prévia autorização da pessoa que à mesma é submetida.

O consentimento, ponto de rutura transformador da relação médico-paciente do tipo vertical numa relação democrática de carácter horizontal, transformou o doente num consumidor de cuidados de saúde[3], cuja dignidade, personalidade, autonomia e autodeterminação assumem posição de primazia. A este cabe decidir sobre a sua saúde e bem-estar físico e psíquico, a partir das suas convicções, objetivos, preferências e valores[4].

Enquanto concretização axiológica do princípio basilar da dignidade humana e simultâneo corolário inelutável dos direitos de personalidade[5], a sua inobservância consubstancia, pois, uma violação do direito fundamental à autodeterminação, independentemente de nos encontrarmos no domínio público ou privado, ou mesmo de existir uma relação de natureza contratual.

---

[2] Guilherme de Oliveira, *Temas de Direito da Medicina*, 2.ª edição (ed.), Coimbra, Coimbra Editora, 2005, p. 110.

[3] André Gonçalo Dias Pereira, *O consentimento informado na Relação Médico-Paciente: Estudo de Direito Civil*, Faculdade de Direito da Universidade de Coimbra, Centro de Direito Biomédico, Coimbra, Coimbra Editora, 2004, páginas (pp.) 20 e 29.

[4] Com esta orientação *vide*, entre tantos, Anabela Susana de Sousa Gonçalves; Bruna de Sousa; Diana Coutinho, *Direito Médico: Responsabilidade Médica, Reprodução Humana Assistida*, Coimbra, Gestlegal, 2024, p. 124, e João Vaz Rodrigues, *O Consentimento informado para o ato médico no ordenamento jurídico português*, Coimbra, Coimbra Editora, 2001, p. 17.

[5] *Vide* André Gonçalo Dias Pereira, *Responsabilidade médica e consentimento informado. Ónus da prova e nexo de causalidade*, s.p., 2009, disponível em https://estudogeral.uc.pt/handle/10316/10577 [consultado em 19/07/2025], e Bruna de Sousa, *Da Perda de Chance na Responsabilidade Civil do Médico*, Coimbra, Almedina, 2022, p. 49.

Sob uma perspetiva histórica e contextual, é usualmente apontado o caso *Schoendorff v. Society of New York Hospitals*, decidido em Nova Iorque, em 1914[6], como a primeira manifestação jurisprudencial em que o consentimento se firmou, de modo inequívoco, como expressão jurídica do direito à autodeterminação do paciente. A situação em apreço reportava-se à realização de uma intervenção cirúrgica para a remoção de um tumor abdominal, não obstante o paciente ter manifestado prévia e expressamente a sua oposição a tal procedimento. Este quadro factual conduziu o Juiz Benjamin Cardozo à formulação de uma máxima que viria a marcar não apenas a evolução doutrinal, como a orientação jurisprudencial subsequente: "Every human being of adult years and sound mind has a right to determine what shall be done with his own body"[7]/[8].

Note-se, porém, que a figura do consentimento não se apresenta como realidade estanque e imutável, antes se tendo revelado alvo de paulatina evolução. Se, por um lado, o período iluminista representou alavanca incomensurável de afirmação de tutela dos direitos humanos, foi precisamente na fase de rescaldo da Segunda Guerra Mundial que se apuraram paradigmáticas alterações, com o surgimento do Código de Nuremberga. Enquanto primeira resposta ético-normativa frente às intervenções médicas não consentidas, este diploma

---

[6] Note-se que há Autores que situam a primeira decisão jurisprudencial no caso *Slater versus Baker & Stapleton*, decidido em 1767, na Inglaterra. Tal posição é, todavia, contestada pela doutrina maioritária que considera a referência ao consentimento, dada a época em que surgiu, mero fruto do acaso. Com esta orientação, João Vaz Rodrigues, *O Consentimento informado para o ato médico no ordenamento jurídico português*, *op. cit*., p. 17.

[7] Cfr. Ruth R. Faden; Tom Beauchamp, *A history and theory of informed consent*, Oxford, Oxford University Press, 1986, p. 123, e Eduardo Dantas; Rafaella Nagaroli, «Consentimento Informado do Paciente Frente às Novas Tecnologias da Saúde (Telemedicina, Cirurgia Robótica e Inteligência Artificial)», *Lex Medicinae – Revista Portuguesa de Direito da Saúde*, Ano 17, número (n.º) 33, 2020, p. 34 (pp. 25-64), disponível em https://www.centrodedireitobiomedico.org/publicacao/lex-medicinae--revista-portuguesa-de-direito-da-saudeano-17n33janeiro-junho-2020/59/ [consultado em 19/07/2025].

[8] A jurisprudência europeia, veículo de afirmação do consentimento informado, revela-se vasta. A título exemplificativo, veja-se, em 1894, o caso decidido pelo *Reichgericht*, no âmbito do qual o pé de uma criança foi amputado de molde a remover um abscesso tuberculoso, quando o progenitor da mesma se tinha oposto especificamente a este tratamento. No que concerne à jurisprudência portuguesa, a consagração desta figura revelou-se tardia, tendo o Supremo Tribunal de Justiça (STJ) proferido a primeira decisão de condenação por violação do consentimento do paciente, a 2 de junho de 2015, no âmbito do processo n.º 1263/06.3TVPRT.P1.S1. Neste sentido, André Dias Pereira, «A consagração do Direito ao Consentimento Informado na Jurisprudência Portuguesa Recente», in *Direito da Saúde: Estudos em Homenagem ao Prof. Doutor Guilherme de Oliveira*, Vol. 3, Coimbra, Almedina, 2016, p. 166.

conformou-se enquanto fonte de inspiração dos mais diversos códigos deontológicos e legislações nacionais sobre a temática em apreço[9].

Entre nós, o consentimento encontra assento vasto e transversal[10].

Iniciando pelo Direito Constitucional, destaca-se, como não podia deixar de ser, o artigo 1.º da Constituição da República Portuguesa (CRP). Consagrante do princípio da dignidade da pessoa humana, projeta-se como postulado em que a pessoa surge como limite e fundamento, fim em si mesma e nunca mero e simples meio[11]. A partir desta premissa, e como afloramentos imediatos, salientam-se, no mesmo diploma, os artigos 24.º a 27.º, bem como o 64.º, enquanto consagrantes do direito à integridade de física e moral, ao livre desenvolvimento da personalidade, à liberdade, à vida e à saúde.

No domínio do Direito Civil, como "afirmação positiva da individualidade do ser humano e da sua conformação livre e consciente"[12], avulta o direito geral de personalidade previsto no artigo 70.º do Código Civil (CC) e respetivas limitações voluntárias, entre as quais assume particular relevo o consentimento, regulado nos artigos 81.º e 340.º do diploma.

Já em sede de Direito Penal, o consentimento surge configurado como causa de exclusão de ilicitude, conforme decorre dos artigos 31.º, 38.º e 149.º do Código Penal (CP). Acresce que, no capítulo respeitante aos crimes contra a liberdade pessoal, os artigos 156.º e 157.º tipificam as intervenções e tratamentos médico-cirúrgicos arbitrários, estabelecendo simultânea e rigorosamente o

---

[9] Rosa Carolina Marques Gaspar Simões, «A Responsabilidade Civil Médica por violação de Consentimento Informado: Novos Rumos do Requisito Causal», *Revista de Direito da Responsabilidade*, Ano 6, 2024, p. 1251 (pp. 1248-1285), disponível em https://revistadireitoresponsabilidade.pt/2024/a-responsabilidade-civil-medica-por-violacao-de-consentimento-informado-novos-rumos-do-requisito-causal-rosa-carolina-marques-gaspar-simoes/ [consultado em 26/08/2025].

[10] Além fronteiras, o consentimento detém assento em diversos instrumentos internacionais. A título exemplificativo, veja-se o artigo 7.º do Pacto Internacional de Direitos Cíveis e Políticos, primeira norma internacional com carácter vinculativo a tutelar o consentimento informado; o artigo 5.º da Convenção sobre os Direitos do Homem e a Biomedicina; o artigo 3.º da Carta dos Direitos Fundamentais da União Europeia (UE); a Diretiva 2005/28/CE da Comissão, de 8 de abril de 2005, que estabelece princípios e diretrizes pormenorizadas de boas práticas clínicas no que respeita aos medicamentos experimentais para uso humano, bem como os requisitos aplicáveis às autorizações de fabrico ou de importação desses produtos; a Declaração Internacional sobre os Direitos Genéticos Humanos; e a Declaração Universal sobre Bioética e Direitos Humanos, ambas da UNESCO.

[11] Seguindo este entendimento, entre muitos outros, Gomes Canotilho; Vital Moreira, *Constituição da República Portuguesa Anotada*, Vol. 1, 4.ª ed., Coimbra, Coimbra Editora, 2007, p. 198, e Gomes Canotilho, *Direito Constitucional e Teoria da Constituição*, 7.ª ed., Coimbra, Almedina, 2003, p. 225.

[12] Paulo Mota Pinto, «O Direito ao Livre Desenvolvimento da Personalidade», in *Portugal-Brasil ano 2000: tema direito*, Coimbra, Coimbra Editora, 1999, p. 166.

dever de esclarecimento do paciente. Igual menção merece o artigo 150.º, na medida em que estabelece que as intervenções e tratamentos médico-cirúrgicos com finalidades terapêuticas não configuram ofensa à integridade física, quando, verificada aquela finalidade, estas sejam resultantes de indicação médica e se verifique qualificação técnica do profissional de saúde.

Finalmente, no plano deontológico, os artigos 6.º e 121.º do Código Deontológico da Ordem dos Médicos (CDOM)[13] estabelecem princípios estruturantes destinados a assegurar que o médico exerce a sua profissão com independência técnica e deontológica, todavia, rapidamente equilibra tais propósitos com reiteradas referências à autonomia do doente e à proteção da sua integridade física e moral. Destaca-se, nesse âmbito, a exigência do consentimento do paciente prevista no artigo 20.º, bem como o dever profissional de esclarecimento consagrado nos artigos 19.º e seguintes[14].

2.1. *Os requisitos de validade do consentimento*

Independentemente da finalidade da intervenção médica – seja ela terapêutica, diagnóstica ou de outra natureza – para que o consentimento se repute válido e eficaz, não basta a sua mera existência formal. Com efeito, além de se revelar pessoal – prestado por aquele que se vai submeter à intervenção –, prévio – concedido antecipadamente à intervenção, garantindo-se um período de reflexão por parte do paciente –, e atual – mantido durante toda a prestação de cuidados de saúde, sem prejuízo da sua revogabilidade, a todo o tempo, nos termos do n.º 2 do artigo 81.º do CC –, o consentimento deverá apresentar-se capaz, livre e esclarecido[15].

Iniciando pela capacidade, exige-se ao paciente o discernimento necessário à compreensão da natureza, alcance e implicações da intervenção, em contrabalança com os seus interesses. O conceito de capacidade para consentir não encontra, portanto, total coincidência com o conceito legal de capacidade negocial consagrada no CC, sendo, antes, uma aferição casuística da maturidade, discernimento e autonomia intelectual do paciente, em face da situação concreta. Quando inexista ou se encontre diminuída será suprida por via da representação

---

[13] Regulamento n.º 707/2016, de 21 de julho, Regulamento de Deontologia Médica.

[14] Em termos de legislação avulsa, destacamos, nunca com carácter taxativo, a Base 2 da Lei de Bases da Saúde (Lei n.º 95/2019, de 4 de setembro); os artigos 3.º e 7.º da Lei n.º 15/2014, de 21 de setembro; a Lei da Investigação Clínica (Lei n.º 21/2014, de 16 de abril); e a Lei sobre a Informação Genética Pessoal e Informação de Saúde (Lei n.º 12/2005, de 26 de janeiro).

[15] Sobre estas características, *vide* Anabela Susana de Sousa Gonçalves; Bruna de Sousa; Diana Coutinho, *Direito Médico: Responsabilidade Médica, Reprodução Humana Assistida*, *op. cit.*, p. 124.

legal, seja através do exercício das responsabilidades parentais, no caso de menores, seja mediante o regime do maior acompanhado, no caso de maiores. Destaca-se, contudo, que o paciente legalmente incapaz poderá, em determinadas circunstâncias, dispor de discernimento bastante para consentir validamente, ao passo que um paciente legalmente capaz poderá, em concreto, não reunir condições volitivas ou intelectuais para tanto[16]/[17].

No que concerne à voluntariedade, o consentimento deverá reportar-se livre de qualquer coação, pressão indevida, ou manipulação externa. A sua essência reside numa manifestação espontânea de vontade, expurgada de qualquer vício que possa afetar a liberdade[18] e que, independentemente de ser manifestada de forma escrita ou verbal, se deverá revelar expressa[19].

Por último, e com especial destaque face ao nosso objeto de estudo, como teremos oportunidade de discorrer *infra*, o consentimento deverá, ainda, revelar-se informado e esclarecido.

2.2. *A dinâmica do consentimento informado*

A doutrina do consentimento tem vindo a ser desenvolvida nos últimos 100 anos, situando-se as primeiras ações de responsabilidade civil com base na ausência de consentimento, como tivemos a oportunidade de verificar *supra*, tanto na Europa como nos Estados Unidos da América (EUA), em meados do século XX. Não obstante, apenas nos últimos 50 anos se firmou a importância

---

[16] Heinrich Ewald Höster; Eva Sónia Moreira da Silva, *A Parte Geral do Código Civil Português*, 2.ª ed., Coimbra, Almedina, 2022, p. 384.

[17] Sobre o requisito da capacidade, *vide*, entre tantos, Anabela Susana de Sousa Gonçalves; Bruna de Sousa; Diana Coutinho, *Direito Médico: Responsabilidade Médica, Reprodução Humana Assistida*, *op. cit.*, pp. 142 e seguintes (ss.); André Gonçalo Dias Pereira, *O consentimento informado na Relação Médico-Paciente: Estudo de Direito Civil*, *op. cit.*, pp. 84 e ss.; e Eduardo Dantas, «Dilemas bioéticos na literacia em saúde: consentimento informado, o exercício da autonomia, e o analfabetismo funcional», *Lex Medicinae – Revista Portuguesa de Direito da Saúde*, Ano 21, n.º 41, 2024, pp. 52 e ss. (pp. 41-55), disponível em https://www.centrodedireitobiomedico.org/publicacao/dilemas-bioeticos-na-literacia-em-saude-consentimento-informado-o-exercicio-da-a/98/ [consultado em 29/08/2025].

[18] O paciente não deverá estar sob o efeito de álcool, drogas, medicamentos, depressão ou qualquer outra substância ou circunstância que, de alguma forma, adulterem a sua vontade.

[19] Com efeito, quanto à forma para consentir, vigora o princípio da liberdade de forma prescrito no artigo 219.º do CC. Não obstante, por imposições de certeza e segurança jurídicas, dá-se preferência à forma escrita, vigorando, entre nós, legislações que, em casos particulares, exigem aquela forma. Veja-se o artigo 14.º, n.º 2, da Lei n.º 32/2006, de 26 de julho, a propósito da procriação medicamente assistida e a alínea l) do artigo 2.º da Lei n.º 21/2014, de 16 de abril, relativa à Investigação Clínica. Sobre estes casos, ainda que excecionais, vastos, *vide*, Anabela Susana de Sousa Gonçalves; Bruna de Sousa; Diana Coutinho, *Direito Médico: Responsabilidade Médica, Reprodução Humana Assistida*, *op. cit*., pp. 126 e ss.

jurídica do direito à informação do paciente, remontando a expressão "*informed consent*" a uma decisão proferida em 1957 pelo Tribunal da Califórnia[20].

Dizer que o consentimento é informado é dizer que o paciente, leigo, quando consentiu na intervenção médica, se encontrava na posse de todas as informações necessárias a um juízo de contraposição da intervenção que lhe foi imposta de forma heterónoma frente ao seu sistema de valores.

De modo instrumental àquele direito encontra-se, concomitantemente, o correspondente dever do médico de informar e esclarecer. Sobre o prestador dos cuidados de saúde, enquanto profissional com competências técnicas, recai o dever de fornecer ao paciente um conjunto de informações imprescindíveis acerca do seu diagnóstico e tratamento, a efeitos de o capacitar de um "poder de decisão consciente"[21]. Donde, o consentimento do paciente nunca será uma decisão solitária, mas partilhada. Ainda que a decisão final lhe caiba unicamente – a este cabe consentir, recusar ou escolher uma intervenção alternativa, ainda que não a mais indicada pelo prestador de cuidados de saúde –, sempre será fruto e reflexo de um diálogo, informações e esclarecimentos prévios[22].

O dever de informar encontra previsão em diversas normas do direito português. Às de aplicação transversal, entre as quais se destacam os artigos 157.º do CP, o 19.º do CDOM e a alínea e) do artigo 2.º da Lei de Bases da Saúde[23], assomam-se outras tantas de aplicação a domínicos específicos, pelo que, na senda de André Dias Pereira, afirmamos existir "uma aparente falta de coerência sistemática" no que concerne às formulações legais deste dever[24]. Não obstante, realizada uma compilação das mesmas, entre os elementos a transmitir encontram-se o diagnóstico e estado de saúde do paciente; natureza e evolução da doença; prognóstico; meios e fins do tratamento; natureza, consequências secundárias, riscos e benefícios do tratamento proposto; alternativas ao tratamento proposto, bem como riscos e consequências do mesmo; riscos e consequências da recusa; e, por último, os aspetos económicos do tratamento.

---

[20] André Dias Pereira, *Responsabilidade médica e consentimento informado. Ónus da prova e nexo de causalidade*, *op. cit.*, s.p., e Anabela Susana de Sousa Gonçalves; Bruna de Sousa; Diana Coutinho, *Direito Médico: Responsabilidade Médica, Reprodução Humana Assistida*, *op. cit*., pp. 126 e ss.

[21] Rosa Carolina Marques Gaspar Simões, «A Responsabilidade Civil Médica por violação de Consentimento Informado: Novos Rumos do Requisito Causal», *op. cit.*, p. 1253.

[22] João Vaz Rodrigues, *O Consentimento informado para o ato médico no ordenamento jurídico português*, *op. cit*., pp. 18 e 224-225.

[23] Lei n.º 95/2019, de 4 de setembro.

[24] André Gonçalo Dias Pereira, *Responsabilidade médica e consentimento informado. Ónus da prova e nexo de causalidade*, *op. cit.*, s.p.

Denote-se que tais formulações, não constituindo um elenco fechado ou taxativo, apontam tão-somente critérios base, *standards* que orientam e elucidam o médico na concretização do dever de informar e esclarecer[25]. O *quantum* informativo associado não se encontra determinado no ordenamento normativo[26] em favor do propósito que as informações transmitidas visam: o esclarecimento do doente. Ora, este há de ser encontrado, necessariamente, de modo casuístico, frente à situação concreta daquele paciente e não outro[27]. Enfim, variando o *quantum* de informação em função de um conjunto de critérios distintos – a natureza do ato, a finalidade curativa ou não curativa prosseguida, a frequência e probabilidade de ocorrência de um determinado risco, entre outros aspetos concretamente pertinentes –, referências genéricas, não especificadas ou atualizadas aos aspetos relevantes *in casu*, não serão suficientes[28]. O consentimento genérico deverá dar assento a um consentimento concreto e individualizado.

Havendo violação do dever de esclarecimento, o consentimento tem-se por ineficaz e, bem assim, a intervenção como ilícita[29].

---

[25] *Ibidem.*

[26] Álvaro da Cunha Gomes Rodrigues, «Consentimento Informado – Pedra Angular da Responsabilidade Criminal do Médico», in Guilherme de Oliveira (coordenação) (coord.), *Direito da Medicina – I*, Coimbra, Coimbra Editora, 1999, p. 40 (pp. 5-51).

[27] Neste sentido, o n.º 5 do artigo 19.º do CDOM.

[28] Rute Teixeira Pedro, «Breves reflexões sobre a reparação de danos causados na prestação de cuidados de saúde com utilização de robots», in Eva Sónia Moreira da Silva; Pedro Miguel Freitas (coord.), *Inteligência artificial e robótica: desafios para o direito do século XXI*, Coimbra, Gestelegal, 2022, p. 164 (pp. 151-186).

[29] Tenha-se presente que o consentimento, em casos devidamente contextualizados, poderá sofrer limitações. Tal verificar-se-á, desde logo, em situações de intervenções urgentes e necessárias em que a prestação de informações não se coaduna com a necessidade urgente da intervenção; nos casos comummente designados de direito a não saber, em que o próprio paciente, ao abrigo do princípio da autonomia e da dignidade da pessoa humana, recusa a transmissão de informações, renunciando ao direito à informação de que é titular; e, ainda, as situações excecionais de privilégio terapêutico, em que a revelação de informações colocaria a vida ou saúde do paciente em níveis extremos de risco. Sobre estes casos, de forma pormenorizada, *vide*, entre tantos, Anabela Susana de Sousa Gonçalves; Bruna de Sousa; Diana Coutinho, *Direito Médico: Responsabilidade Médica, Reprodução Humana Assistida*, *op. cit*., pp. 151 e ss., e André Gonçalo Dias Pereira, *O consentimento informado na Relação Médico-Paciente: Estudo de Direito Civil*, *op. cit.*, pp. 286 e ss.

3. A IA na prestação de cuidados de saúde

Parafraseando Ana Elizabete Ferreira e André Dias Pereira, "há muito que a relação médico-paciente não é uma relação a dois"[30]. Destarte, desde o século passado que o setor médico tem vindo a atravessar uma linha evolutiva de digitalização crescente. O recurso gradual à telemedicina[31] – desde a teleconsulta à teleintervenção cirúrgica –, a digitalização do arquivo dos ficheiros clínicos[32], assim como a digitalização da prescrição de fármacos[33] são concretizações e repercussões claras da acentuação da digitalização e interconectividade que a vida moderna enfrenta.

Não obstante, poderá dizer-se que a revolução que temos vindo a presenciar atingiu o seu auge. Seja com uma presença irreverente, seja de modo mais comezinho – ainda que nunca despiciendo –, as máquinas sofisticadas, dotadas de autonomia e cujo propósito último se firma na emulação da inteligência humana, invadem hoje os nossos hospitais. Sistemas de captação e análise de imagens médicas como o *DeepMind*[34] ou o *Convolutional CapsNet*, utilizado, com grande potencial, em plena pandemia de Covid-19[35]; sistemas que apresen-

---

[30] Ana Elisabete Ferreira; André Dias Pereira, «Uma ética para a medicina pós-humana: propostas ético-jurídicas para a mediação das relações entre humanos e robôs na saúde», in Nelson Rosenvald; Joyceane Bezerra de Menezes; Luciana Dadalto (coord.), *Responsabilidade Civil e Medicina*, São Paulo, Editora Focus, 2020, p. 2 (pp. 1-19), disponível em https://www.researchgate.net/publication/344950550_Uma_Etica_para_a_Medicina_Pos-Humana [consultado em 18/08/2025].

[31] A regulação desta modalidade pode ser encontrada nos artigos 46.º e ss. do CDOM.

[32] Sobre os registos eletrónicos de saúde e os dados de saúde eletrónicos, importa destacar o recente Regulamento (UE) 2025/327 do Parlamento Europeu e do Conselho, de 11 de fevereiro de 2025, relativo ao Espaço Europeu de Dados de Saúde e que altera a Diretiva 2011/24/UE e o Regulamento (UE) 2024/2847.

[33] O Despacho n.º 7979P/2015, de 20 de julho, do INFARMED, na sequência do Decreto-Lei n.º 106A/2010, de 1 de outubro, tornou obrigatória a prescrição eletrónica médica – banalmente denominada *receita sem papel* – no Serviço Nacional de Saúde.

[34] Sistema de IA de análise de imagens oftálmicas que visa o diagnóstico e a avaliação do progresso de patologias várias, bem como a análise de imagens de células cancerígenas. Cfr. André Gonçalo Dias Pereira, «Inteligência Artificial, Saúde e Direito: Considerações Jurídicas em torno da Medicina de Conforto e da Medicina Transparente», *Revista Julgar*, n.º 45, 2021, p. 240 (pp. 235 a 261), disponível em https://julgar.pt/inteligencia-artificial-saude-e-direito-consideracoes-juridicas-em-torno-da-medicina-de-conforto-e-da-medicina-transparente/ [consultado em 25/08/2025].

[35] Guilherme Pires Silva, «Autoria e imputação na Responsabilidade Civil Médica pelo uso de sistemas de diagnóstico dotados de Inteligência Artificial: O futuro não começa amanhã», *Revista de Direito da Responsabilidade*, Ano 6, 2024, p. 235 (pp. 233-269), disponível em https://revistadireitoresponsabilidade.pt/2024/autoria-e-imputacao-na-responsabilidade-civil-medica-pelo-uso- de-sistemas-de-diagnostico-dotados-de-inteligencia-artificial-guilherme-pires-da-silva/ [consultado em 23/03/2025].

tam opções de tratamento por ordem de pertinência e probabilidade de sucesso como o *Watson for Oncology* da IBM[36]; sistemas de avaliação da probabilidade de desenvolvimento de patologias de esquizofrenia, diabetes e cancro, como o *Deep Patient*[37]; sistemas multifacetados como o *Dr. A.I.* da empresa *HelthTap* que não só verifica sintomas dos pacientes, como os orienta e recomenda cuidados médicos personalizados[38]; sistemas de distribuição de material médico e recolha de exames como o *robot Moxi*[39]; *robots* cirúrgicos; próteses inteligentes; sistemas de monitorização de parâmetros fisiológicos[40]; ou sistemas de diagnóstico com capacidades para produzir relatórios provenientes da análise de imagens de raio-X como o *Oxipit*[41], são meras indicações dentro de uma panóplia de exemplos ínfima de como a IA – aliada a outras técnicas e tecnologias como a robótica, a nanotecnologia e a genómica – mergulha em contexto sanitário[42].

Concretamente em Portugal, existem já diversas iniciativas que visam promover uma saúde digital, mormente, pela transformação da IA de uma promessa de futuro para uma realidade presente. *Ad exemplum*, indique-se o projeto *Derm-AI*, que, detendo como principal objetivo a análise de lesões cutâneas e outras condições de pele, propõe um apoio ao diagnóstico e tratamento de doenças dermatológicas, classificando ainda os casos de prioridade, como os de

---

[36] Marta Susana Lobo, «Responsabilidade Médica e Inteligência Artificial», *Lex Medicinae – Revista Portuguesa de Direito da Saúde*, Ano 20, n.º 39, 2023, p. 82 (pp. 73-91), disponível em https://www.centrodedireitobiomedico.org/publicacao/revista-portuguesa-de-direito-da-saudelex-medicinaeano- 20n-39janeiro-junho-2023/66/ [consultado em 25/03/2025].

[37] *Ibidem*, p. 83.

[38] André Gonçalo Dias Pereira, «Inteligência Artificial, Saúde e Direito: Considerações Jurídicas em torno da Medicina de Conforto e da Medicina Transparente», *op. cit.*, p. 235.

[39] Eduardo Dantas; Rafaella Nagaroli, «Consentimento Informado do Paciente Frente às Novas Tecnologias da Saúde (Telemedicina, Cirurgia Robótica e Inteligência Artificial)», *op. cit.*, pp. 53-54.

[40] Rute Teixeira Pedro, «Breves reflexões sobre a reparação de danos causados na prestação de cuidados de saúde com utilização de robots», *op. cit.*, p. 156.

[41] Repare-se que este sistema de IA, inclusive, já recebeu a marcação CE. Cfr. Marta Paixão Cansado, *Inteligência Artificial no setor da saúde: desafios jurídicos e regulação*, GEE Papers, 187, Gabinete de Estratégia e Estudos do Ministério da Economia, 2024, p. 5, disponível em https://www.gee.gov.pt/pt/estudos-e-seminarios/gee-papers-category/33911-inteligencia-artificial-no-setor-da-saude-desafios-juridicos-e-regulacao [consultado em 10/08/2025], e Mindy Nunez Duffoure; Dominick S. Giovanniello, «The Autonomous AI Physician: Medical Ethics and Legal Liability», in AA. VV., *Law, Governance and Technology Series: Multidisciplinary Perspectives on Artificial Intelligence and the Law*, Vol. 58, Cham, Springer, 2024, p. 208 (pp. 207-228).

[42] Com uma descrição pormenorizada de como a IA pode surgir no setor da saúde, *vide* o documento produzido pela Organização Mundial de Saúde, *Ethics and governance of artificial intelligence for health: WHO guidance*, World Health Organization, 2021, pp. 6-16, disponível em www.who.int/publications/i/item/9789240029200 [consultado em 01/07/2025].

suspeita de cancro da pele, bem como o projeto *SCOUT.AI*, desenvolvido para o serviço de triagem, aconselhamento e encaminhamento do SNS 24, por meio de processamento de linguagem natural[43].

Efetivamente, e porque tal exercício se revelaria tarefa herculana destinada ao fracasso, não é nosso propósito levar aqui a cabo uma enunciação exaustiva dos inesgotáveis modos de utilização que esta disruptiva tecnologia mobiliza. Todavia, este pequeno trecho deixa clara a revolução a que temos presenciado no universo clínico, impedindo o seu branqueamento ou negação[44]. Os singulares progressos verificados ao nível da ciência e da tecnologia biomédica, bem como a sua integração no contexto médico, têm sido, aliás, destacados pelas alterações que se perspetivam na dinâmica médico-paciente e das quais resultam vantagens várias para a comunidade, no seu todo, bem como para cada paciente considerado de modo individual.

Afastam-se as tarefas anteriormente rotineiras, repetitivas e burocráticas, abrindo-se margem de tempo e energia para a investigação; otimizam-se os recursos disponíveis; garantem-se as condições necessárias para diagnósticos prematuros, com custos mais baixos; reduz-se o número de procedimentos invasivos e, consequentemente, o impacto dos tratamentos; permitem-se recuperações mais rápidas; um maior rigor e precisão dos atos cirúrgicos; o aumento da fiabilidade dos resultados[45]; estabelecem-se os meios e as condições para o abandono da típica pontualidade do acompanhamento médico, o qual se passa a desenvolver de forma mais próxima e contínua[46]; e, como zénite da questão,

---

[43] Sobre estes projetos e os restantes em curso na Unidade de *Advanced Analytics*, Inteligência Artificial e Robótica (UIA), *vide* Serviços Partilhados do Ministério da Saúde (SPMS), EPE, *Inteligência Artificial na Saúde em Portugal: Regulamentação, Impactos e Perspetivas de Futuro*, White Paper, 2025, pp. 4-5, disponível em https://www.spms.min-saude.pt/2025/03/inteligencia-artificial-na-saude-em-portugal-regulamentacao-impactos-e-perspetivas-de-futuro/ [consultado em 20/08/2025].

[44] André Gonçalo Dias Pereira compara a prestação de cuidados de saúde ao paradigma verificado nas agências de viagem. Tal como nas últimas os bilhetes de avião e os programas de turismo são, agora, na grande maioria, comprados por via da *internet*, dispensando a presença de um ser humano, também no futuro "uma aplicação ou um site especializado, na área médica, poderá fazer-nos prescindir de consultas presenciais ou mesmo por telemedicina" – *vide*, «Inteligência Artificial, Saúde e Direito: Considerações Jurídicas em torno da Medicina de Conforto e da Medicina Transparente», *op. cit*., p. 237.

[45] Rute Teixeira Pedro, «Breves reflexões sobre a reparação de danos causados na prestação de cuidados de saúde com utilização de robots», *op. cit*., p. 157.

[46] Rafaella Nogaroli; José Luiz de Moura Feleiros Júnior, «Ethical Challenges of Artificial Intelligence in Medicine and the Triple Semantic Dimensions of Algorithmic Opacity with Its Repercussions to Patient Consent and Medical Liability», in AA. VV., *Law, Governance and Technology Series: Multidisciplinary Perspectives on Artificial Intelligence and the Law*, Vol. 58, Cham, Springer, 2024, p. 230 (pp. 229-250), e Marta Paixão Cansado, *Inteligência Artificial no setor da saúde: desafios jurídicos e regulação*, *op. cit.*, p. 5.

institui-se uma medicina personalizada, preditiva, preventiva e proativa, ou seja, uma medicina que deixa de concentrar a sua atenção no tratamento da patologia para deter como foco a antecipação do seu diagnóstico e prevenção[47].

O objetivo é claro e simples: acelerar o processo de diagnóstico, reduzir o risco de erro humano e aumentar a qualidade e a esperança de vida[48].

3.1. *Os desafios da IA no plano jurídico*

É usual apontar-se a década de 50 como aquela em que surgiram os primeiros sistemas de IA. À data, estes reportavam-se – ainda que de modo verdadeiramente inovador para a época – como máquinas simples. Operando através da manipulação de símbolos e regras básicas previamente disponibilizadas de *se isto, então aquilo*, estes sistemas detinham como propósito único o encontro de uma solução específica, para um problema que concretamente lhes era colocado.

O paradigma contemporâneo da IA, apresenta-se, todavia, substancialmente distinto. Longe da rigidez dos sistemas simbólicos, os sistemas de IA contemporâneos, baseados em modelos de *machine* e *deep learning*, caracterizam-se pela aprendizagem automática. Os algoritmos não se limitam a seguir regras prévias e estritamente fornecidas pelo ser humano, passando a aprender, modificar e adaptar o seu comportamento com base nos dados que autonomamente reco-

---

[47] Comummente denominada medicina dos 4 P's, a partir da análise pormenorizada e sequenciação do fenótipo e do genótipo de cada indivíduo, esta define a estratégia terapêutica apropriada, bem como o momento adequado para a introduzir e, ainda, as doenças concretas a que cada indivíduo, de modo individualizado, está predisposto a verificar, associando-lhes a correspondente forma de prevenção ou minoria. Pelo contrário, quando a patologia já se verifique, este tipo de medicina permite uma adequação e otimização da terapêutica medicamentosa a aplicar. A propósito da medicina personalizada, *vide* Eduardo Figueiredo, *Direito da Nanobiotecnociência*, Coimbra, Almedina, 2021, pp. 62 e ss.; Ana Elisabete Ferreira; André Dias Pereira, «Uma ética para a medicina pós-humana: propostas ético-jurídicas para a mediação das relações entre humanos e robôs na saúde», *op. cit.*, p. 15 (pp. 1-19), e Miguel Kfouri Neto; Rafaella Nogaroli, «O consentimento do paciente no admirável mundo novo de robôs de assistência de inteligência artificial para diagnóstico médico», in Rodrigo da Guia Silva; Gustavo Tepedino (coord.), *O Direito Civil na era da inteligência artificial*, 1.ª ed., São Paulo, Thomson Reuters Brasil, 2020, p. 140 (pp. 139-164), disponível em https://www.academia.edu/44127111/O_consentimento_do_paciente_no_admir%C3%A1vel_mundo_novo_de_rob%C3%B4s_de_assist%C3%AAncia_%C3%A0_sa%C3%BAde_e_algoritmos_de_intelig%C3%AAncia_artificial_para_diagn%C3%B3stico_m%C3%A9dico [consultado em 18/08/2025]. Concretamente sobre os riscos e desvantagens desta modalidade, *vide* André Gonçalo Dias Pereira, «Inteligência Artificial, Saúde e Direito: Considerações Jurídicas em torno da Medicina de Conforto e da Medicina Transparente», *op. cit.*, pp. 238 e 245-248.

[48] Com orientação semelhante, *vide* os pontos 33 e 34 da Resolução do Parlamento Europeu, de 16 de fevereiro de 2017, que contém recomendações à Comissão sobre disposições de Direito Civil sobre Robótica (2015/2103(INL)), disponível em https://op.europa.eu/pt/publication-detail/-/publication/13fd56d0-8a65-11e8-ac6a-01aa75ed71a1 [consultado em 17/04/2025].

lhem, alcançando soluções que resultam do seu próprio traquejo e experiência arrecadada[49].

É precisamente nesta capacidade adaptativa que reside a difusão da IA, tanto na sociedade em geral, como, de modo particularmente significativo, na seara da prestação de cuidados de saúde. Não obstante, "[w]ith great innovations comes great challenges"[50]. Se a IA vem preencher grande parte da lacuna existente entre o desempenho humano e a perfeição[51], permitindo até figurar que, dentro em breve, a figura do médico se torne obsoleta, num olhar mais atento e cuidado, abundam desafios, riscos e preocupações cuja debelação cumpre ao Direito acautelar.

Quid iuris se a máquina algorítmica falha num diagnóstico; se o algoritmo utilizado pelo sistema de IA é instruído com dados discriminatórios; se se verifica uma falha da *internet* durante uma cirurgia robótica; se o *robot* assistente deixa cair, sem qualquer razão justificativa e aparatosamente o paciente que transportava; ou mesmo se o *robot* cirurgião, sob a mácula da autonomia, decide atuar de forma autónoma e imprevisível provocando danos no paciente ou, *in extremis*, a morte?

São precisamente estas inquietações que se encontram patentes no Livro Branco sobre a IA da Comissão Europeia. Sistematizando os principais riscos em torno do desenvolvimento e aplicação da tecnologia, destacam-se os erros de classificação e previsão; os riscos de viés nos dados de treino, validação e implementação; a ausência de transparência no desenho, desenvolvimento e implantação dos sistemas de IA, que impedem a sua rastreabilidade e explicabilidade de resultados; o uso incorreto; as lacunas na atribuição de responsabilidades; a violação do direito à privacidade e proteção de dados; a desconfiança dos clíni-

---

[49] Para uma análise desenvolvida sobre as várias etapas evolutivas da IA, bem como o atual estado da arte, *vide*, entre outros, Stuart Russell; Peter Norving, *Artificial Intelligence: A Modern Approach*, 4.ª ed., Harlow, Pearson, 2022, pp. 35 e ss.; Arlindo Oliveira, *Inteligência Artificial*, Lisboa, Fundação Francisco Manuel dos Santos, 2019; e Arlindo L. Oliveira; Mário A. T. Figueiredo, «Artificial Intelligence: Historical Context and State of the Art», in AA. VV., *Law, Governance and Technology Series: Multidisciplinary Perspectives on Artificial Intelligence and the Law*, Vol. 58, Cham, Springer, 2024, p. 230 (pp. 3-25).

[50] António Cruz Oliveira; Tiago Branco da Costa, «A Saúde Digital: Responsabilidade artificial?», in AA.VV., *As palavras necessárias – Estudos em comemoração dos 30 anos da Escola de Direito por ocasião do centenário de Francisco Salgado Zenha*, Vol. 1, Braga, UMinho Editora, 2023, p. 197 (pp.193-225).

[51] Mindy Nunez Duffoure; Dominick S. Giovanniello, «The Autonomous AI Physician: Medical Ethics and Legal Liability», *op. cit*., p. 210.

cos; bem como a necessidade da tecnologia utilizada por parte dos profissionais de saúde se revele, garantidamente, confiável e transparente[52].

Não menos relevante, a par da natural carência de empatia e compaixão – dimensões naturais da prática da medicina, que a fazem transcender as ciências puramente formalistas, geridas pela estatística e pela matemática[53] –, a ausência de *feedback háptico*, isto é, a ausência da informação transmitida pelo toque ou informação cinestésica oferecida pelos tecidos, que permite identificar alterações na sua consistência e elasticidade, um maior cuidado na sua correta e adequada manipulação, bem como o teste do grau de dor e resistência do paciente[54].

Nesta seara, embora se proclame que "a mão robótica/informática/artificial ultrapassou a mão humana e/ou melhorou a *performance* humana face ao seu equivalente biológico médico"[55], alcançando resultados mais céleres, eficientes e precisos, tratamos de uma tecnologia que não é absolutamente precisa, motivo pelo qual, sem nos delongarmos mais, se coloca a questão de saber, enfim, como afirmar o âmbito do dever de informar e esclarecer inerente ao consentimento informado, frente a uma utilização crescente de tecnologias inteligentes?

3.2. *IA e consentimento informado: contornos específicos*

Atendendo a tudo o que até então já foi dito, não nos parecem existir dúvidas de que o recurso a sistemas de IA constituirá, precisamente, um dos elementos a integrar o acervo de informação transmitida ao paciente[56]. A atual

---

[52] Comissão Europeia, *Livro Branco sobre a inteligência artificial: Uma abordagem europeia virada para a excelência e a confiança*, de 19 de fevereiro de 2020 [COM (2020) 65 final], p. 28, disponível em https://eurlex.europa.eu/legalcontent/PT/TXT/PDF/?uri=CELEX:52020IP0276&from=PT [consultado em 08/09/2025].

[53] Mindy Nunez Duffoure; Dominick S. Giovanniello, «The Autonomous AI Physician: Medical Ethics and Legal Liability», *op. cit*., p. 210.

[54] Ana Elisabete Ferreira; André Dias Pereira, «Uma ética para a medicina pós-humana: propostas ético-jurídicas para a mediação das relações entre humanos e robôs na saúde», *op. cit*., p. 2.

[55] *Ibidem*, p. 3.

[56] Com a mesma posição, Rute Teixeira Pedro, «Breves reflexões sobre a reparação de danos causados na prestação de cuidados de saúde com utilização de robots», *op. cit.*, p. 156; Ana Elisabete Ferreira; André Dias Pereira, «Uma ética para a medicina pós-humana: propostas ético-jurídicas para a mediação das relações entre humanos e robôs na saúde», *op. cit*., p. 7; e o Parlamento Europeu, quando, no ponto 13 da Resolução de 16 de fevereiro de 2017 que estabelece recomendações à Comissão Europeia a propósito de normas de direito civil sobre a robótica (2015/2103(INL)), disponível em https://op.europa.eu/pt/publication-detail/-/publication/13fd56d0-8a65-11e8-ac6a-01aa75ed71a1 [consultado em 01/09/2025], indica, precisamente, que o consentimento "deve ser solicitado e obtido antes de qualquer interação homem-máquina".

doutrina do consentimento informado detém, dentro de objetivos vários, que o paciente "compreenda todas as circunstâncias do tratamento proposto e da tecnologia utilizada para esse fim, bem como as alternativas terapêuticas, possibilitando a escolha informada"[57], pelo que a recusa do tratamento levado a cabo por um *robot* configura "reivindicação basilar da transição tecnológica na saúde"[58].

O propósito deverá ser, sempre, garantir a autodeterminação da pessoa e, para tal, dar ao paciente a possibilidade de escolha, num primeiro momento, quanto à realização do procedimento com ou sem a intervenção da IA e, num segundo momento – ainda que, à data, longínquo –, quando a utilização da IA não figure como alternativa, mas, dada a sua proliferação, o padrão único de atendimento, dar ao doente a possibilidade de escolha de se submeter ou não à intervenção[59].

Por outro lado, atendendo a que, numa situação típica, ausente da intervenção de um sistema de IA, ao médico cabe não só informar da doença como explicá-la, não só informar do tratamento, como explicá-lo em linguagem menos técnica e compreensível; no caso de uma intervenção cirúrgica deter a assistência de um *robot* inteligente, não só a sua utilização deverá ser transmitida, como o papel específico do sistema no procedimento deverá ser informado[60], *v.g.*, o momento em que este irá intervir e os moldes da intervenção. Isto porque, partindo do princípio de que o paciente sabe que um sistema de IA se irá imiscuir na intervenção, a sua decisão poderá ser distinta caso o sistema detenha como função única a monitorização dos seus parâmetros fisiológicos ou, pelo invés, ser o próprio sistema, de forma autónoma, a operar, reduzindo-se a presença do médico à supervisão da atuação.

Devendo a informação transmitida caracterizar-se pela completude e pela lealdade, sopesa que também os riscos decorrentes do recurso à IA e as vantagens da mesma derivadas – principalmente em comparação com os riscos e

---

[57] Eduardo Dantas; Rafaella Nagaroli, «Consentimento Informado do Paciente Frente às Novas Tecnologias da Saúde (Telemedicina, Cirurgia Robótica e Inteligência Artificial)», *op. cit*., p. 37.

[58] Ana Elisabete Ferreira; André Dias Pereira, «Uma ética para a medicina pós-humana: propostas ético-jurídicas para a mediação das relações entre humanos e robôs na saúde», *op. cit*., p. 13.

[59] Neste sentido, Glenn Cohen, «Informed Consent and Medical Artificial Intelligence: What to Tell the Patient?», *op. cit*., p. 1462; Rute Teixeira Pedro, «Breves reflexões sobre a reparação de danos causados na prestação de cuidados de saúde com utilização de robots», *op. cit*., pp. 164-165; e Márcia Santana Fernandes; José Roberto Goldin, «Artificial Intelligence and Decision Making in Health: Risks and Opportunities», in AA.VV., *Law, Governance and Technology Series: Multidisciplinary Perspectives on Artificial Intelligence and the Law*, Vol. 58, Cham, Springer, 2024, p. 194 (pp. 187-206).

[60] Glenn Cohen, «Informed Consent and Medical Artificial Intelligence: What to Tell the Patient?», *op. cit.*, p. 1439.

vantagens da não utilização de tal tecnologia – deverão ser transmitidos[61]. Já não merecendo acolhimento a ética médica defensora do encobrimento dos riscos com potencial desencorajador da intervenção, o paciente deverá ser informado, além dos riscos gerais, limitações, benefícios e alternativas associadas ao procedimento, da possibilidade de falha técnica, eletricidade, conexão ou ataque de terceiro; dos riscos derivados da necessidade de implementação de um plano de contingência para conversão do procedimento para a sua versão tradicional e consequências daí derivadas, *v.g.*, maior tempo sob o efeito da anestesia ou a necessidade de uma incisão maior[62]; e, ainda, os níveis de treino e experiência do médico em causa com o sistema que se propõe utilizar, bem como os inerentes riscos de erro médico derivados do deficiente ou insuficiente treino[63].

De qualquer forma, não nos parece ser exigível ao prestador de cuidados de saúde uma explicação de modo completo das informações específicas de funcionamento do sistema na sua componente de *hardware* e *software* e, assim, as formas concretas com que a IA chegou ou chegará a determina conclusão ou atuação[64].

Efetivamente, o conteúdo do dever de informar e esclarecer tem sido causa de esforços, quer na doutrina, quer na jurisprudência, havendo, ainda, dúvidas quando ao seu concreto conteúdo[65]. Não obstante, é hoje assente que o médico não tem a obrigação de transmitir ao doente todas e quaisquer informa-

---

[61] Com a mesma orientação, ainda que a propósito de *robots* inteligentes, Rute Teixeira Pedro, «Breves reflexões sobre a reparação de danos causados na prestação de cuidados de saúde com utilização de robots», *op. cit.*, p. 164.

[62] Como exemplo elucidativo deste tipo de ocorrência, cita-se um incidente registado na cidade de Filadélfia, durante uma intervenção cirúrgica assistida pelo *robot Da Vinci*. No decurso do procedimento, o equipamento começou a apresentar mensagens de erro, pelo que, e face à ineficácia das intervenções por parte da equipa de suporte técnico, se tornou imperativo converter a cirurgia robótica numa abordagem cirúrgica convencional. Entre o momento em que a equipa médica resolveu não utilizar o sistema de IA e o momento em que os equipamentos necessários à continuação da cirurgia foram colocados a pontos, passaram cerca de 45 minutos. Cfr. Eduardo Dantas; Rafaella Nagaroli, «Consentimento Informado do Paciente Frente às Novas Tecnologias da Saúde (Telemedicina, Cirurgia Robótica e Inteligência Artificial)», *op. cit*., p. 41.

[63] Seguimos de perto Michael J. Rigby, «Ethical Dimensions of Using Artificial Intelligence in Health Care», *AMA Journal of Ethics*, n.º 2, Vol. 21, 2019, p. 122 (pp. 121-124), disponível em https://journalofethics.ama-assn.org/article/ethical-dimensions-using-artificial-intelligence-health-care/2019-02 [consultado em 21/08/2025], e Rafaella Nogaroli; José Luiz de Moura Feleiros Júnior, «Ethical Challenges of Artificial Intelligence in Medicine and the Triple Semantic Dimensions of Algorithmic Opacity with Its Repercussions to Patient Consent and Medical Liability», *op. cit.*, p. 244.

[64] No mesmo sentido, Rute Teixeira Pedro, «Breves reflexões sobre a reparação de danos causados na prestação de cuidados de saúde com utilização de robots», *op. cit.*, p. 165.

[65] André Gonçalo Dias Pereira, *Responsabilidade médica e consentimento informado: Ónus da prova e nexo de causalidade*, *op. cit.*, s.p.

ções, ao mais ínfimo pormenor, mas tão-só a informação necessária, pertinente e suficiente para que o paciente entenda realmente a situação em causa[66]. Aliás, a factualidade de o dever de informar se concretizar muito exigente tem vastas desvantagens: os custos são mais elevados, o profissional está a despender tempo com a transmissão de informações que, porventura, não relevam na decisão do doente, e, no limite, o paciente, abalroado com informação excessiva, paradoxalmente, não conseguirá, a final, distinguir os riscos significativos dos riscos insignificantes[67].

Note-se que, inclusive, quando o tratamento proposto é determinado inteiramente pelo prestador dos cuidados de saúde, o mesmo não tem de explicar em que informações específicas é que a escolha se baseou, *v.g.*, as memórias vagas de uma palestra sobre a temática patológica *in casu* a que assistiu na faculdade, a prática levada a cabo pelos médicos com os quais laborou durante o período de residência, os resultados que obteve da sua experiência médica, ou o artigo que leu recentemente sobre as mais recentes práticas médicas[68].

Acresce que será comum o profissional de saúde não deter conhecimento ao mais ínfimo pormenor sob o modo de funcionamento do sistema. Tratamos de tecnologia que se pauta e caracteriza pela complexidade, exigindo conhecimentos em áreas do saber que ultrapassam aquelas em que tipicamente os prestadores de cuidados de saúde atuam e dominam[69]. No mais, importa sublinhar que, se, de uma banda, o paciente tem o direito a ser informado numa linguagem leiga, acessível e que atenda à sua capacidade cognitiva, ao seu nível intelectual e cultural, de outra banda, a informação que equacionamos mostra-se de complexidade proporcional à tecnologia de onde deriva, sendo entre nós assente a dificuldade de comunicar informações complexas com precisão aos pacientes.

Face ao problema, recorrendo às soluções preconizados pelo recente Regulamento da IA[70], verificamos que, estando em causa sistemas que atuam no domínio da saúde, se esses sistemas não constituírem uma prática inaceitável nos

---

[66] Anabela Susana de Sousa Gonçalves; Bruna de Sousa; Diana Coutinho, *Direito Médico: Responsabilidade Médica, Reprodução Humana Assistida*, *op. cit*., p. 146.

[67] André Gonçalo Dias Pereira, *Responsabilidade médica e consentimento informado: Ónus da prova e nexo de causalidade*, *op. cit*., s.p.

[68] Glenn Cohen, «Informed Consent and Medical Artificial Intelligence: What to Tell the Patient?», *op. cit.*, p. 1457.

[69] Rute Teixeira Pedro, «Breves reflexões sobre a reparação de danos causados na prestação de cuidados de saúde com utilização de robots», *op. cit.*, p. 165.

[70] 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 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.

termos do artigo 5.º do Regulamento[71] – estando garantido que, por um lado, os seus resultados não figuram condicionantes aos juízos e decisões clínicas, nem como obstáculo à efetividade de um consentimento livre, autónomo e consciente e, por outro lado, que a sua utilização não explore a vulnerabilidade típica das pessoas sujeitas a cuidados de saúde[72] –, serão considerados, nos termos do disposto no n.º 2 do artigo 6.º e nas alíneas a) e d) do ponto 5 do anexo III, como sistemas de IA de risco elevado[73].

Classificados enquanto tal, estes sistemas distinguem-se pela vasta panóplia de deveres por que se fazem acompanhar e cujo escopo primordial consiste na prevenção dos danos que, pela sua natureza e elevada probabilidade de impacto, podem causar à saúde, à segurança e aos direitos fundamentais dos cidadãos. Desenvolvidos nos artigos 9.º a 19.º do Regulamento, entre o mencionado rol de obrigações, assumem especial relevância as obrigações de autonomia e explicabilidade e, conotada a esta última e dela pressuposto necessário, a transparência.

Efetivamente, o Regulamento não só estabelece o respeito pela autonomia humana como princípio orientador, como o eleva a imperativo ético diretor da atuação dos prestadores de cuidados de saúde, assim garantindo que a tecnologia de que se cuida seja desenvolvida, implementada e utilizada de forma confiável. Ora, como pressuposto da autodeterminação daqueles que com ela interagem, surge o requisito da supervisão humana. Previsto no artigo 14.º, este torna clara a visão da IA, não como substituta do ser humano, mas como ferramenta ao seu serviço. Neste sentido, a sua atuação e resultados estarão, sempre, dependentes de prévia validação humana, garantindo-se, para uma supervisão e controlo adequados – por meio de ferramentas interface homem-máquina – que o responsável pela implantação detenha capacidade para controlar o sistema, interpretar os seus resultados e intervir quando tal se revele necessário. Como sublinha André

[71] Ao longo dos seus treze capítulos e anexos, com o objetivo de melhorar o funcionamento do mercado interno mediante a previsão de um regime jurídico uniforme, o Regulamento promove uma estrutura de riscos de ordem decrescente, pelo que nele encontramos sistemas de riscos inaceitáveis, associados a práticas de IA proibidas [capítulo II], sistemas de IA de risco elevado [capítulo III]; sistemas de IA de risco limitado [capítulo IV]; e sistemas de IA de risco mínimo [capítulo IV].

[72] Cfr. o artigo 5.º, n.º 1, alíneas a) e b), do Regulamento da IA.

[73] Para uma análise sobre o Regulamento da IA, *vide* Dário Moura Vicente, «Inteligência Artificial e Iniciativas Internacionais», in Manuel Lopes Rocha; Rui Soares Pereira (coord.); Ana Coimbra Trigo (colab.), *Inteligência Artificial & Direito*, Coimbra, Almedina, 2020, pp. 93-105; e José Ricardo Marcondes Ramos, «Supervisão, classificação e certificação dos sistemas de IA na Proposta de Regulamento sobre Inteligência Artificial», in Susana Aires de Sousa (coord.), *Direito em Mudança – A Proposta de Regulamento Europeu sobre Inteligência Artificial – Algumas questões jurídicas*, Coimbra, Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2023, pp. 31-64.

Dias Pereira, "perante uma IA de risco elevado, acrescenta-se uma IH (Inteligência Humana)"[74].

Em conotação, e para concretização dessa supervisão, torna-se imprescindível assegurar um grau suficiente de transparência no funcionamento dos sistemas[75]. Afastando o problema das *black boxes*, o Regulamento impõe que o processo decisório do sistema e os critérios subjacentes à sua atuação sejam suficientemente transparentes. Só deste modo, e na posse de tais conhecimentos, o responsável pela implantação poderá decidir quando vai utilizar o sistema de IA, se o estiver a utilizar, parar tal utilização, e decidir – porque os consegue interpretar – se os resultados obtidos podem ou não ser utilizados. É a rastreabilidade e a explicabilidade que estão na base.

Ora, a transposição destes ensinamentos para o contexto da prestação de cuidados de saúde exige, assim, que o paciente esteja esclarecido de que a tecnologia tem um grau de falibilidade e que, no final, sempre caberá ao médico, porque na posse dos conhecimentos necessários para tal, decidir sobre a aplicação dos resultados por aquela apresentados.

4. Conclusão

No seio do Admirável Mundo Novo, o fenómeno disruptivo, inteligente e digital em curso vaticina uma transformação da sociedade que se perspetiva eterna. No universo clínico, a introdução crescente de tecnologia inteligente, embora portadora de inegáveis e heterógenos benefícios, erige, indubitavelmente, desafios éticos e jurídicos de elevada complexidade. Convidando à reflexão sobre as questões que interpelam, especial destaque assume a necessidade de compatibilização entre o uso de sistemas de IA e o respeito pela autonomia e autodeterminação do doente, corolários do instituto do consentimento informado.

Regressamos ao ponto de partida com as interrogações colocadas a título introdutório: terá o prestador dos cuidados de saúde de informar o doente da presença, recurso e utilização de um sistema de IA? Em caso afirmativo, exigir-se-á igualmente uma explicação da forma de funcionamento geral do sistema, ou dos parâmetros em que concretamente se baseou a decisão?

O tema incorpora questões outrora levantadas por Ulrick Beck quanto ao debate sobre o impacto da tecnologia na vida dos seres humanos: como que-

[74] André Gonçalo Dias Pereira, «Inteligência Artificial, Saúde e Direito: Considerações Jurídicas em torno da Medicina de Conforto e da Medicina Transparente», *op. cit*., p. 250.

[75] Artigo 13.º do Regulamento da IA.

remos viver? O que há de humano no ser humano e de natural na natureza que precisa de ser protegido?[76]

Quanto ao que nos apraz, parece-nos clara a resposta positiva frente àquela primeira questão. Havendo a equação da possibilidade de a prestação de cuidados de saúde ser levada a cabo com recurso à utilização de um sistema de IA, este é um dos aspetos – de importância capital, necessária e obrigatória, afirme-se – que compõe o núcleo duro do acervo de informações a transmitir. O consentimento pressupõe um processo de escolha, livre, esclarecida e informada, pelo que a discussão das possibilidades terapêuticas configura não mero ornamento fáctico da relação entre o médico e o paciente, mas elemento constitutivo da eficácia do consentimento prestado.

Todavia, entendemos não se revelar necessário que tal informação abranja uma explicação específica e completa sobre o funcionamento do sistema de IA empregue, nem a lógica subjacente a determinado resultado, por irrelevante à tomada de decisão do paciente. A *ratio* em que reside o consentimento habita no conhecimento dos riscos, das vantagens e consequências da intervenção – nomeadamente, e por comparação à realização da intervenção sem a IA – e não na apreensão dos processos lógico-matemáticos subjacentes à tecnologia que a coadjuva. Essa exigência de compreensão técnica será exigida, outrossim, ao prestador de cuidados de saúde, como modo de garantia de cumprimento do requisito da supervisão a que a abordagem normativa do Regulamento da IA sujeita os sistemas inteligentes que atuam em contexto médico, enquanto sistemas de risco elevado.

Urge, pois, melhorar a formação dos profissionais de saúde que recorrem à IA. Não basta que a IA seja desenvolvida de modo explicável se o profissional que a utiliza não é portador das bases necessárias à compreensão dessa explicabilidade. Garantindo-se a formação, garante-se que o prestador de cuidados de saúde é detentor do conhecimento necessário para a transmissão das informações e explicações necessárias ao consentimento, bem como à supervisão da atuação do sistema, afastando as situações de confiança cega nos resultados por ele apresentados.

Em jeito de conclusão, à luz de tudo quanto foi dito, porque o desenvolvimento tecnológico é uma manifestação da humanidade, impõe-se afirmar que a introdução da IA na prestação de cuidados de saúde não poderá constituir mecanismo de supressão ou instrumento que esvazia a autonomia e autodeterminação do paciente. Ao invés, sendo circunstância que as afirma e intensifica, deverá ser redobrada a acuidade na concretização do direito ao consentimento informado, contudo, sem nunca o desvirtuar.

---

[76] Ulrich Beck, *La Sociedad del Riesgo: Hacia una nueva modernidad*, Barcelona, Buenos Aires, México, Paidós Básica, 1998, p. 34.

DECRETO-LEI N.º 2/2025, DE 23 DE JANEIRO – O REGIME SANCIONATÓRIO APLICÁVEL À VIOLAÇÃO DAS OBRIGAÇÕES IMPOSTAS PELO REGULAMENTO EUROPEU DE GOVERNAÇÃO DE DADOS

*Ana Sirage Coimbra**

Resumo: O Regulamento Governação de Dados estabelece que os Estados-Membros adotem todas as medidas necessárias para garantir a sua aplicação, devendo prever sanções efetivas, proporcionadas e dissuasivas. Este artigo apresenta os aspetos gerais do regime contraordenacional do Decreto-Lei n.º 2/2025, de 23 de janeiro, que executa o Regulamento (UE) 2022/868, do Parlamento Europeu e do Conselho, de 30 de maio de 2022, relativo à governação europeia de dados.

Palavras-chave: Regulamento Governação de Dados; Contraordenações; Procedimento.

* Jurista na Administração Pública e Investigadora não Integrada no Centro de Investigação em Justiça e Governação (JusGov) da Escola de Direito da Universidade do Minho. https://orcid.org/0009-0000-5368-1733. E-mail: ac8649@gmail.com.

Abstract: The Data Governance Regulation establishes that Member States shall take all necessary measures to ensure its application, and must provide for effective, proportionate, and dissuasive sanctions. This article presents the general aspects of the administrative offenses regime of Decree-Law No. 2/2025, of January 23th, which implements Regulation (EU) 2022/868 of the European Parliament and of the Council, of 30th of May 2022, on European data governance.

Keywords: Data Governance Regulation; Administrative offenses; Procedure.

1. Introdução

Tendo em vista apoiar os cidadãos e as empresas da União Europeia (UE), o Regulamento (UE) 2022/868 do Parlamento Europeu e do Conselho, de 30 de maio de 2022, relativo à governação europeia de dados [doravante, Regulamento Governação de Dados (RGD)], pretende promover a disponibilização de mais informação para reutilização e simplificar a partilha de dados em diversas áreas (como, por exemplo, na saúde, no ambiente, no setor energético, na agricultura, nas finanças, na indústria transformadora e na administração pública). O Considerando (3) do RGD assume a necessidade de aperfeiçoamento das condições de partilha de dados no mercado interno, "criando um quadro harmonizado para o intercâmbio de dados e estabelecendo determinados requisitos básicos para a governação de dados, prestando especial atenção à facilitação da cooperação entre os Estados-Membros". Neste âmbito, o RGD define as condições para a reutilização de determinados dados detidos por organismos do setor público que têm na sua posse um amplo repositório de dados protegidos por direitos de terceiros cuja utilização no formato dados abertos não é permitida (tais como, propriedade intelectual, segredos comerciais ou dados pessoais), mas que podem ser reutilizados ao abrigo de regras específicas nacionais ou da UE. Nos casos em que essa reutilização é admitida, os organismos do setor público devem respeitar as condições de reutilização fixadas no RGD. Salvaguarda-se, porém, a exclusão do âmbito de aplicação do RGD da informação detida pelo setor público relacionada com a segurança pública, a defesa e a segurança nacional dos Estados-Membros, bem como dados provenientes de procedimentos de contratação pública abrangidos pelo âmbito de aplicação da Diretiva 2009/81/CE do Parlamento Europeu e do Conselho.

Assim, do RGD decorre a obrigação de cada Estado-Membro designar um ou vários organismos competentes, em determinadas áreas, para apoiar os organismos do setor público que, nesse quadro jurídico, concedem ou recusam acesso, para fins de reutilização, a certas categorias de dados protegidos. Considera-se "Reutilização"[1] a utilização, por pessoas singulares ou coletivas, de dados detidos por organismos do setor público, realizada para fins comerciais ou não comerciais que não correspondem à finalidade inicial da missão de serviço público para a qual os dados foram produzidos, excetuando o intercâmbio de

[1] Encontramos idêntica definição na alínea g) do n.º 1 do artigo 3.º da Lei n.º 26/2016, de 22 de agosto, na redação atual (que regula o acesso aos documentos administrativos e à informação administrativa, incluindo em matéria ambiental, transpondo para a ordem jurídica interna a Diretiva 2003/4/CE, do Parlamento Europeu e do Conselho, de 28 de janeiro de 2003, relativa ao acesso do público às informações sobre ambiente).

dados entre organismos do setor público exclusivamente no desempenho das suas missões de serviço público[2] – artigo 2.º, ponto 2), do RGD.

O RGD também procede à regulação dos prestadores de serviços de intermediação de dados, que nos termos do ponto 11) do artigo 2.º do RGD, são "serviços que visam estabelecer relações comerciais para efeitos de partilha de dados entre um número indeterminado de titulares dos dados e detentores dos dados, por um lado, e utilizadores de dados, por outro, através de meios técnicos, jurídicos ou outros, inclusive para o exercício dos direitos dos titulares dos dados em relação aos dados pessoais"[3]. Na prática, estes operadores são intermediários neutrais que asseguram a relação entre indivíduos e empresas que detêm dados e entidades que queiram utilizar esses dados. As exigências desses serviços visam garantir que aqueles prestadores funcionam como mediadores de confiança na partilha de dados.

O RGD cria ainda um enquadramento legal para o altruísmo de dados, que, de acordo com o ponto 16) do artigo 2.º do RGD, "abrange a partilha voluntária de dados, com base no consentimento dos titulares dos dados para o tratamento dos respetivos dados pessoais ou na autorização, por parte de outros detentores dos dados, da utilização dos seus dados não pessoais, sem que esses titulares ou detentores procurem ou recebam uma gratificação que vá além de uma compensação pelos custos em que incorrem ao disponibilizarem os seus dados, para fins de interesse geral, previstos no direito nacional, se aplicável, tais como os cuidados de saúde, a luta contra as alterações climáticas, a melhoria da mobilidade, a facilitação do desenvolvimento, produção e divulgação de estatísticas oficiais, a melhoria da prestação dos serviços públicos, a elaboração de políticas públicas ou a investigação científica de interesse geral".

Assim, cabe a cada Estado-Membro designar uma ou mais autoridades competentes em matéria de serviços de intermediação de dados, bem como uma

---

[2] Nos termos do n.º 4 do artigo 19.º da Lei n.º 26/2016, de 22 de agosto, na redação atual, a troca de documentos administrativos entre os órgãos e entidades referidos no artigo 4.º, exclusivamente no âmbito do desempenho das suas funções e dos fins de interesse público que lhes compete prosseguir, não constitui reutilização.

[3] As exclusões constam das alíneas a) a d) do ponto 11 do artigo 2.º do RGD, respetivamente: "a) Serviços que obtêm dados junto dos detentores dos dados e agregam, enriquecem ou transformam os dados obtidos com o objetivo de lhes acrescentar um valor substancial e licenciam a utilização dos dados resultantes aos utilizadores de dados, sem estabelecer uma relação comercial entre os detentores dos dados e os utilizadores dos dados; b) Serviços centrados na intermediação de conteúdos protegidos por direitos de autor; c) Serviços exclusivamente utilizados por um único detentor dos dados para permitir a utilização dos dados detidos por esse detentor dos dados, ou utilizados por várias pessoas coletivas no seio de um grupo fechado, inclusive no âmbito de relações com fornecedores ou clientes ou colaborações contratualmente estabelecidas, em especial os que tenham como principal objetivo assegurar funcionalidades de objetos e dispositivos ligados à Internet das coisas; d) Serviços de partilha de dados oferecidos por organismos do setor público que não visam estabelecer relações comerciais".

ou mais autoridades competentes em matéria de registo das organizações de altruísmo de dados. O RGD impõe igualmente a cada Estado-Membro a obrigação de estabelecer um ponto de informação único que funcione como interface para os utilizadores que pretendam reutilizar categorias específicas de dados detidos por organismos do setor público, e que disponibilize, por via eletrónica, uma lista pesquisável de recursos que preveja uma panorâmica de todos os recursos de dados disponíveis.

De notar que tanto os dados pessoais como os não pessoais são abrangidos pelo RGD, ainda que os considerandos[4] iniciais ressalvem que este regulamento "não deverá ser interpretado como criando uma nova base jurídica para o tratamento de dados pessoais no quadro de quaisquer das atividades regulamentadas, nem como alterando os requisitos em matéria de informação previstos no Regulamento (UE) 2016/679". O mesmo considerando acrescenta ainda que, havendo conflito entre o RGD e o direito da UE em matéria de proteção de dados pessoais ou o direito nacional adotado em conformidade com esse direito da União, deverá prevalecer o direito da União ou o nacional aplicável em matéria de proteção de dados pessoais.

O Decreto-Lei n.º 2/2025, de 23 de janeiro (doravante DLERGD, isto é, decreto-lei de execução do Regulamento Governação de Dados), veio executar o RGD, designando o organismo competente para apoiar os organismos do setor público, o ponto de informação único, as autoridades competentes em matéria de serviços de intermediação de dados, bem como a autoridade competente em matéria de registo das organizações de altruísmo de dados, e entrou em vigor em 28 de janeiro de 2025.

2. O regime sancionatório do Decreto-Lei n.º 2/2025, de 23 de janeiro, que executa o Regulamento (UE) 2022/868 do Parlamento Europeu e do Conselho, de 30 de maio de 2022

O artigo 34.º do Regulamento Governação de Dados determina que compete aos Estados-Membros definir as regras relativas às sanções aplicáveis em caso de violação das obrigações referentes às transferências de dados não pessoais para países terceiros e da obrigação de notificação aplicável aos prestadores de serviços de intermediação de dados, bem como quando se verifique o incumprimento das condições para a prestação de serviços de intermediação de dados e das condições para o registo como organização de altruísmo de dados reconhecida. Neste sentido, o DLERGD estabelece o regime sancionatório resultante do disposto no RGD.

[4] Considerando (4) do Regulamento (UE) 2022/868.

À semelhança do que sucede com outros diplomas específicos [como, por exemplo, o Código da Estrada (artigo 136.º), o Código do Trabalho (artigo 553.º), a Lei-Quadro das Contraordenações Ambientais (artigo 21.º) ou o Regime Jurídico das Contraordenações Económicas (artigo 17.º)], no DLERGD, as contraordenações classificam-se como leves, graves e muito graves, tendo em conta a gravidade e relevância dos interesses violados (artigo 6.º do DLERGD). A este propósito, e conforme se pronunciou o Acórdão do Supremo Tribunal de Justiça n.º 6/2018[5], de 26/09/2018, Processo n.º 215/15.7T8ACB.C1 -A. S1, Diário da República, 1.ª série, n.º 219, de 14/11/2018, a gravidade de uma infração "é determinada pela gravidade da ilicitude pressuposta pelo legislador. Ao classificar uma dada infração como grave o legislador considerou-a, em abstrato, portadora de uma ilicitude considerável, o que terá desde logo determinado uma moldura da coima com limites mínimos e/ou máximos superiores àqueles que foram determinados para as contraordenações que entendeu como sendo de gravidade menor ou de média gravidade. Depois, em função do caso concreto, e dentro dos limites da coima prevista pelo legislador, ir-se-á determinar a medida concreta da sanção em atenção às finalidades de punição das coimas e em atenção à culpa do agente. Todos estes elementos poderão ser determinantes para que se entenda que, pese embora se trate de uma contraordenação grave, portadora de uma ilicitude, em abstrato, grave, atento o caso concreto dever-se-á entender que o agente deverá ser punido com uma sanção próxima do seu limite mínimo. Porém, não se pode considerar que, atento o caso concreto, a ilicitude da conduta diminua de gravidade, depois de o legislador a ter classificado como sendo uma contraordenação grave, porque portadora de uma ilicitude considerada grave. Na verdade, sempre que o legislador, de forma geral e abstrata, classifica a infração como sendo grave, não poderá o julgador modificar a lei atribuindo menor gravidade àquela ilicitude".

Na tipificação dos ilícitos, são diferenciados dois grupos contraordenacionais. O primeiro grupo abrange as contraordenações no âmbito da prestação de serviços de intermediação de dados (artigo 7.º do DLERGD) e o segundo as contraordenações no âmbito do altruísmo de dados (artigo 8.º do DLERGD).

[5] Disponível em https://diariodarepublica.pt/dr/detalhe/acordao-supremo-tribunal-justica/6-2018-116967080 [acedido em 02/09/2025].

Tabela 1

Contraordenações no Decreto-Lei n.º 2/2025, de 23 de janeiro

Grupo 1		**Grupo 2**
Artigo 7.º Intermediação de Dados		Artigo 8.º Altruísmo de Dados
Muito graves	Alíneas a) a f) do n.º 1 do artigo 7.º	Alíneas a) a g) do n.º 1 do artigo 8.º
Graves	Alíneas a) a k) do n.º 2 do artigo 7.º	Alíneas a) a e) do n.º 2 do artigo 8.º
Leves	Alíneas a) a e) do n.º 3 do artigo 7.º	Não aplicável

Quanto à aplicação no espaço, o artigo 12.º do DLERGD estabelece que salvo disposição em contrário em tratado ou convenção internacional, aquele decreto-lei é aplicável aos factos praticados em território português, independentemente da nacionalidade do agente. De forma análoga ao que sucede no artigo 4.º do Regime Geral das Contraordenações (Decreto-Lei n.º 433/82, de 27 de outubro, na redação atual), esta norma consagra como princípio geral o *princípio da territorialidade*, ou seja, este decreto-lei é aplicável aos factos praticados no território nacional, ressalvando, porém, os tratados ou convenções em contrário. Isto significa, de acordo com Simas Santos e Lopes de Sousa[6], "que a lei contra-ordenacional é ainda aplicável a quaisquer factos cometidos fora do território português e que este esteja obrigado a julgar por convenção internacional ou acordo no âmbito da cooperação judiciária, e deixará de ser aplicável quando, pelos mesmos meios, Portugal tenha aceitado a aplicação de uma outra lei contra-ordenacional".

Relativamente ao montante das coimas aplicáveis, estão previstos no artigo 10.º do DLERGD, e também variam consoante o grau de gravidade da infração. Acresce ainda a diferenciação resultante da qualidade do agente da infração, consoante se trate de pessoa singular ou coletiva. Pelo seu interesse relativamente a esta questão, mencionamos o teor do Acórdão do Tribunal Constitucional n.º

[6] Manuel Simas Santos; Jorge Lopes de Sousa, *Contraordenações – Anotações ao Regime Geral*, 6.ª ed., Lisboa, Áreas Editores, 2011, p. 108.

397/2019, de 04/07/2019, Processo n.º 78/17[7], 1.ª Secção: "o Tribunal Constitucional já entendeu – embora sobre a perspetiva de análise do princípio da igualdade – que a diferença – por vezes significativa – entre os montantes das coimas aplicáveis a pessoas singulares e a pessoas coletivas não violava tal princípio, porque a «radical distinção de natureza entre pessoas singulares e colectivas exclui, desde logo, a existências de igualdade fáctica que constitui o necessário pressuposto para que se possa considerar a operatividade do princípio jurídico-constitucional da igualdade» (Acórdão n.º 569/98). Como se escreveu no Acórdão n.º 110/2012: «(…) o legislador pode instituir tratamento diferenciado em relação a pessoas coletivas com base justamente na específica natureza e características dessas entidades no confronto com as pessoas físicas que detenham personalidade individual. Essa fundamental distinção explica que se tenha assistido no âmbito do direito sancionatório, e em especial no domínio do direito de mera ordenação social, a uma progressiva responsabilização das pessoas coletivas, que se tem caracterizado também pelo estabelecimento de coimas de montantes mais elevados do que os determinados para as pessoas singulares em relação ao mesmo tipo de infração. Nesse sentido, o agravamento da moldura abstrata das coimas aplicáveis às pessoas coletivas foi consagrado como princípio geral no Regime Geral das Contraordenações, como ressalta do seu artigo 17.º, que prevê como montante máximo da coima € 44 891,81 ou € 22 445,91, em caso de negligência, por contraponto aos limites de € 3 740,98 e € 1 870,49, para as pessoas singulares (cfr. Paulo Pinto Albuquerque, *Comentário do Regime Geral das Contraordenações*, Universidade Católica Editora, Lisboa, 2011, págs. 76-77)».

[7] Disponível em https://www.tribunalconstitucional.pt/tc/acordaos/20190397.html [acedido em 02/09/2025].

Tabela 2

Montante das coimas no Decreto-Lei n.º 2/2025, de 23 de janeiro

Classificação	Agente	Moldura do dolo		Moldura da negligência[8]	
Muito graves	Pessoas singulares	Mínimo	Máximo	Mínimo	Máximo
		€ 2.500,00	€ 3.740,00	€ 1.250,00	€ 1.870,00
	Pessoas coletivas	Mínimo	Máximo	Mínimo	Máximo
		€20.000,00	€44.890,00	€10.000,00	€22.445,00
Graves	Pessoas singulares	Mínimo	Máximo	Mínimo	Máximo
		€500,00	€ 2.500,00	€250,00	€ 1.250,00
	Pessoas coletivas	Mínimo	Máximo	Mínimo	Máximo
		€ 5.000,00	€20.000,00	€ 2.500,00	€10.000,00
Leves	Pessoas singulares	Mínimo	Máximo	Mínimo	Máximo
		€100,00	€1.000,00	€50,00	€500,00
	Pessoas coletivas	Mínimo	Máximo	Mínimo	Máximo
		€1.000,00	€ 2.500,00	€500,00	€ 1.250,00

Por força do disposto no n.º 5 do artigo 10.º do DLERGD, às contraordenações muito graves, para além da coima, podem ser aplicadas, em função da gravidade da infração e da culpa do agente, as sanções acessórias que se mostrem adequadas, nos termos da legislação em matéria contraordenacional que é subsidiariamente aplicável[9]. Pelo que, neste diploma, o legislador optou por restringir a aplicação de sanções acessórias apenas às contraordenações muito graves. Para Ferreira Antunes[10], as sanções acessórias, "taxativamente catalogadas, só podem ser determinadas pela lei em função – proporcional e adequada – da gravidade da infração (censurabilidade do facto) e da culpa (censurabilidade do agente). Tais exigências significam, desde logo, que fica excluída a possibilidade de criação de sanções acessórias de aplicabilidade automática".

[8] De acordo com o n.º 4 do artigo 10.º do DLERGD, a negligência é punível, sendo os limites mínimos e máximos das coimas aplicáveis reduzidos a metade.

[9] Artigos 28.º e 29.º do Regime Jurídico das Contraordenações Económicas, *ex vi* artigo 13.º do DLERGD.

[10] Manuel Ferreira Antunes, *Contraordenações e Coimas – Regime Geral*, 2.ª ed., Lisboa, Petrony Editora, 2013, p. 151.

No mesmo sentido, segundo Paulo Albuquerque[11], "[…] O princípio da proporcionalidade aplica-se às sanções acessórias, por força do disposto no artigo 18.º da CRP […] a determinação concreta da medida da sanção acessória deve ser feita em função da gravidade objetiva do facto e da gravidade subjetiva da culpa do agente, do que também decorre que a infração contra-ordenacional tem de ser praticada no exercício ou por causa das atividades com que se relacionam as sanções acessórias, não sendo suficiente que a conduta típica tenha lugar por ocasião do desenvolvimento de tais actividades, mas sem estar conexionada com elas […] O princípio da não automaticidade das sanções acessórias implica que a sua aplicação depende da gravidade da infração e da culpa do agente […]".

A UE reconhece a importância para a sociedade civil na utilização e reutilização de dados gerados ou recolhidos por organismos do setor público ou outras entidades a expensas dos orçamentos públicos[12]. No entanto, existem determinadas categorias de dados que se encontram em bases de dados públicas (como dados comerciais confidenciais, dados que estão sujeitos a confidencialidade estatística e os dados protegidos por direitos de propriedade intelectual de terceiros, incluindo segredos comerciais e dados pessoais) que nem sempre são disponibilizados, por falta de conhecimentos ou de meios adequados para facilitar esse tipo de reutilização (inclusive para atividades de investigação ou de inovação de interesse público, ainda que esse acesso seja possível nos termos do direito da União aplicável). Dada a sensibilidade desses dados, a sua disponibilização impõe a conformidade com determinados requisitos processuais técnicos e jurídicos, quer para assegurar o respeito dos direitos de terceiros sobre os dados visados, quer para evitar efeitos prejudiciais na salvaguarda dos direitos fundamentais, no princípio da não discriminação e na proteção de dados. Segundo Moniz[13], no que diz respeito à estratégia para os dados, assistimos a "um ponto de viragem no trajeto da UE: à necessidade de criar restrições relativamente ao tratamento de dados pessoais, justificadas sobretudo por razões de proteção dos direitos fundamentais, junta-se agora um imperativo – uma quase urgência – de *abertura* e, principalmente, de promoção de soluções jurídicas que suportem a reutilização dos dados em geral, independentemente da sua natureza".

---

[11] Paulo Pinto de Albuquerque, *Comentário do Regime Geral das Contraordenações – à luz da Constituição da República, da Convenção Europeia dos Direitos Humanos e da Carta dos Direitos Fundamentais da União Europeia*, 2.ª ed. atualizada, Lisboa, Universidade Católica Editora, 2022, pp. 128-129.

[12] Diretiva (UE) 2019/1024 do Parlamento Europeu e do Conselho, de 20 de junho de 2019, relativa aos dados abertos e à reutilização de informações do setor público, transposta para a ordem jurídica portuguesa através da Lei n.º 68/2021, de 26 de agosto, que aprova os princípios gerais em matéria de dados abertos.

[13] Graça Canto Moniz, *A Estratégia da UE para os Dados – Que Contributos para a regulação dos dados?*, Coimbra, Edições Almedina S.A., 2024, p. 8.

Relativamente às contraordenações muito graves previstas nas alíneas a) a f) do n.º 1 do artigo 7.º do DLERGD, estas estão relacionadas com a prestação dos serviços de intermediação de dados, em particular com os requisitos exigíveis às entidades que prosseguem estes serviços (designadamente, proibição de práticas restritivas na prestação dos serviços de intermediação dos dados[14], cumprimento de deveres de informação aos detentores dos dados, em caso de transferência, acesso ou utilização não autorizados dos dados não pessoais que tenham partilhado[15] e obrigação de implementação de medidas de segurança e proteção de dados[16]), no sentido de assegurar que a sua atividade é desenvolvida nos termos estabelecidos no RGD. Assinalando a importância da segurança dos dados, Pinto e Filipe[17] realçam que, no contexto atual, "os dados são o motor da inovação tecnológica e das operações empresariais. Desde o comércio eletrónico aos serviços financeiros, saúde e transportes, quase todas as áreas dependem de grandes volumes de dados para funcionarem eficientemente. No entanto, esta dependência crescente traz consigo desafios significativos no que respeita à segurança".

Destacamos, desde logo, a obrigação de notificação que impende sobre os prestadores de serviços de intermediação de dados que tencionem prestar os serviços de intermediação de dados referidos no artigo 10.º à autoridade competente em matéria de serviços de intermediação de dados [cf. alínea a) do n.º 1 do artigo 7.º do DLERGD]. O cumprimento desta obrigação de notificação é uma condição prévia para a prestação do serviço de intermediação de dados (conforme decorre dos n.os 4 e 5 do artigo 11.º do RGD), garantindo que as entidades prestadoras reúnem os requisitos necessários para operar naquela área de atividade. Funciona como uma espécie de acreditação, que permite a identificação das entidades que podem usar o título de "prestador de serviços de intermediação de dados reconhecido na União" na sua comunicação oral e escrita, bem como um logótipo comum (n.º 9 do artigo 11.º do RGD). Por sua

---

[14] As condições comerciais (incluindo o preçário) para a prestação de serviços de intermediação de dados a um detentor ou utilizador dos dados não podem estar dependentes de o detentor dos dados ou o utilizador dos dados utilizar outros serviços prestados pelo mesmo prestador de serviços de intermediação de dados ou por uma entidade com ele relacionada e, caso assim seja, da proporção em que o detentor dos dados ou utilizador dos dados utiliza esses outros serviços – cf. alínea c) do n.º 1 do artigo 7.º do DLERGD.

[15] Cf. alínea e) do n.º 1 do artigo 7.º do DLERGD.

[16] Medidas que assegurem um nível de segurança adequado do armazenamento, do tratamento e da transmissão de dados não pessoais, bem com o mais elevado nível de segurança possível do armazenamento e da transmissão de informações sensíveis ao nível da concorrência – cf. alínea f) do n.º 1 do artigo 7.º do DLERGD.

[17] Bernardo Pinto; Luís Filipe, *O Poder da IA Generativa − Uma nova relação entre a Administração Pública e o Cidadão*, Lisboa, INA Editora, 2025, p. 73.

vez, em linha com o princípio da finalidade do tratamento[18] (as finalidades devem ser *legítimas, determinadas e explícitas*)[19], para além de estar vedada a utilização dos dados relativamente aos quais se presta serviços de intermediação para outros fins que não seja colocá-los à disposição dos utilizadores de dados [cf. alínea b) do n.º 1 do artigo 7.º do DLERGD], esta atividade pressupõe uma autonomia orgânica do ente prestador no exercício da prestação de serviços de intermediação [excluindo os serviços identificados nas alíneas a) a d) do n.º 11 do artigo 2.º do RGD], quando se requer que seja uma pessoa coletiva distinta[20]. Consideramos que esta exigência visa ultrapassar eventuais incompatibilidades associadas à existência de conflitos de interesses na prestação dos serviços de intermediação de dados, nas situações em que os prestadores tivessem de decidir sobre concretas operações de tratamento na qualidade de responsáveis, prejudicando a necessária isenção.

Já quanto às contraordenações graves enunciadas nas alíneas a) a k) do n.º 2 do artigo 7.º do DLERGD, prendem-se, essencialmente, com a violação de formalidades ou procedimentos inerentes ao desenvolvimento da prestação do serviço de intermediação de dados, permitindo à autoridade competente controlar e supervisionar o cumprimento dos requisitos estabelecidos por parte dos prestadores de serviços de intermediação de dados. A nosso ver, visam essencialmente garantir que o funcionamento da atividade após a notificação à autoridade competente se mantém dentro dos parâmetros da regulação. Por um lado, no âmbito dos deveres de informação, obrigam as entidades abrangidas a uma comunicação atualizada de determinados factos, tais como a notificação da cessação da atividade de intermediação de dados [cf. alínea b) do n.º 2 do artigo 7.º do DLERGD], a prestação de informações pelo representante legal do prestador de serviços de intermediação de dados quando solicitado pela autoridade competente [cf. alíneas c) e j) do n.º 2 do artigo 7.º do DLERGD], ou a garantia de uma continuidade aceitável da prestação dos serviços de intermediação de dados, nas situações de insolvência do prestador dos serviços [cf. alínea g) do n.º 2 do artigo 7.º do DLERGD]. Por sua vez, na dimensão do relacionamento com o detentor dos dados ou o titular dos dados, vigoram os princípios da transpa-

[18] Princípio a que também alude o artigo 5.º do Regulamento Geral sobre a Proteção de Dados [Regulamento (UE) 2016/679, de 27 de abril], ao estabelecer que os dados pessoais não podem ser tratados posteriormente de uma forma incompatível com as finalidades do tratamento inicial.

[19] O mesmo princípio é válido para a contraordenação prevista na alínea d) do n.º 1 do artigo 7.º do DLERGD, que estabelece que os dados relativos a qualquer atividade de uma pessoa singular ou coletiva recolhidos para efeitos da prestação do serviço de intermediação de dados apenas podem ser utilizados para o desenvolvimento daquele específico serviço de intermediação de dados.

[20] Sendo que um prestador de serviços de intermediação de dados que não esteja estabelecido na União Europeia, mas que ofereça os serviços de intermediação de dados referidos no artigo 10.º do RGD num dos Estados Membros em que esses serviços são prestados terá de designar um representante legal para o efeito – n.º 3 do artigo 11.º do RGD.

rência e lealdade no tratamento dos dados [quando, por exemplo, a utilização de certos instrumentos e serviços adicionais para facilitar o intercâmbio de dados (como o armazenamento temporário, a curadoria, a conversão, a anonimização e a pseudonimização), depende de pedido ou aprovação expressa por parte do detentor ou do titular dos dados – cf. alínea d) do n.º 2 do artigo 7.º do DLERGD –, ou no dever de esclarecimento prévio do titular dos dados quanto às condições da utilização dos dados – cf. alínea d) do n.º 2 do artigo 7.º do DLERGD], bem como os princípios da justiça, da não discriminação e da proporcionalidade no acesso aos serviços de intermediação de dados, designadamente quanto aos preços e termos do serviço [cf. alínea i) do n.º 2 do artigo 7.º do DLERGD]. Por último, ao nível da segurança e integridade dos dados, assegurar a existência de procedimentos adequados à prevenção de práticas abusivas ou fraudulentas pelas partes interessadas no acesso ao serviço de intermediação de dados [cf. alínea f) do n.º 2 do artigo 7.º do DLERGD], a implementação das medidas técnicas, jurídicas e organizativas apropriadas para impedir a transferência ou o acesso a dados não pessoais que sejam ilegais, nos termos do direito da União Europeia ou do direito nacional [cf. alínea h) do n.º 2 do artigo 7.º do DLERGD], bem como que a transferência internacional ou o acesso governamental de autoridades administrativas a dados não pessoais detidos na União apenas ocorra em determinadas condições [cf. alínea k) do n.º 2 do artigo 7.º do DLERGD]. Importa notar que, segundo Neves[21], com a transformação da organização administrativa e do serviço público, a prossecução de "tarefas e/ou finalidades normativamente reconhecidas como públicas não pressupõe, pois, que sejam levadas a cabo pela Administração Pública. A relação entre a Administração Pública e o serviço público não é uma relação de correspondência. Cabe às autoridades públicas (europeias, estaduais, regionais ou locais, consoante os casos, embora numa lógica integrada e multinível do exercício da função administrativa) esta opção «a definição das obrigações e missões de serviço público», «regular os mercados e… assegurar que os operadores realizem as missões de serviço público que lhe forem confiadas», tudo nos limites do DUE e das políticas públicas europeias".

No caso das contraordenações leves, mencionadas nas alíneas a) a e) do n.º 3 do artigo 7.º do DLERGD, estas visam aspetos operacionais no âmbito da prestação do serviço de intermediação de dados, designadamente a observância do princípio da finalidade quando a prestação de serviço de intermediação de dados pela mesma pessoa coletiva é realizada em conjunto com outras atividades desse prestador de serviços [alínea a) do n.º 3 do artigo 7.º do DLERGD], promovendo a integridade dos suportes de dados [alínea b) do n.º 3 do artigo 7.º do DLERGD], assegurando a interoperabilidade com outros serviços de intermediação de dados através da adoção de normas abertas de uso corrente no

[21] Ana Fernanda Neves, *Direito da Organização Administrativa e Serviço Público*, Coimbra, Edições Almedina S.A., 2023, p. 32.

setor em que os prestadores de serviços de intermediação de dados operam [alínea c) do n.º 3 do artigo 7.º do DLERGD], especificando a jurisdição do país terceiro em que a utilização dos dados se destina a ser efetuada [concedendo aos titulares dos dados instrumentos para dar e retirar o consentimento, e aos detentores dos dados meios para dar e remover a autorização para o tratamento de dados – alínea d) do n.º 3 do artigo 7.º do DLERGD] e mantendo um registo atualizado da atividade de intermediação dos dados [alínea e) do n.º 3 do artigo 7.º do DLERGD].

Agora no âmbito do altruísmo de dados, as contraordenações muito graves estão previstas nas alíneas a) a g) do n.º 1 do artigo 8.º do DELRGD. Neste grupo, as infrações também estão associadas ao incumprimento dos requisitos exigíveis às entidades que prosseguem aquelas atividades [designadamente, das condições exigidas para inscrição num registo público nacional de organizações de altruísmo de dados (cf. alínea a) do n.º 1 do artigo 8.º do DLERGD), das obrigações de notificação à autoridade competente das alterações ao registo (cf. alínea b) do n.º 1 do artigo 8.º do DLERGD), dos requisitos de transparência destas entidades (cf. alínea c) do n.º 1 do artigo 8.º do DLERGD), dos deveres de informação aplicáveis aos titulares dos dados ou detentores dos dados (cf. alínea d) do n.º 1 do artigo 8.º do DLERGD) e da observância do princípio da finalidade do interesse geral na utilização dos dados – cf. alínea e) do n.º 1 do artigo 8.º do DLERGD], incluindo das regras[22] estabelecidas pela Comissão Europeia através de atos delegados, das quais resultem obrigações para as organizações de altruísmo de dados.

Na vertente da relação com os titulares dos dados ou detentores dos dados, deverá ser assegurada a disponibilização de instrumentos para a obtenção do consentimento por parte dos titulares dos dados ou a autorização para o tratamento dos dados disponibilizados pelos detentores dos dados, bem como para a simplificação da retirada desse consentimento ou dessa autorização [cf. alínea f) do n.º 1 do artigo 8.º do DLERGD].

---

[22] De acordo com as alíneas a) a d) do n.º 1 do artigo 22.º do RGD: "a) Requisitos de informação adequados para assegurar que sejam facultadas aos titulares dos dados e aos detentores dos dados, antes de estes darem consentimento ou autorização para o altruísmo de dados, informações suficientemente pormenorizadas, claras e transparentes sobre a utilização dos dados, os instrumentos para dar e retirar o consentimento ou autorização e as medidas tomadas para evitar a utilização abusiva dos dados partilhados com a organização de altruísmo de dados; b) Requisitos técnicos e de segurança adequados para garantir um nível de segurança adequado do armazenamento e do tratamento dos dados, bem como dos instrumentos para dar e retirar o consentimento ou a autorização; c) Roteiros de comunicação com uma abordagem multidisciplinar com vista a sensibilizar as partes interessadas pertinentes, em especial os detentores dos dados e os titulares dos dados que possam eventualmente partilhar os seus dados, para o altruísmo de dados, para a designação como «organização de altruísmo de dados reconhecida na União» e para o conjunto de regras; d) Recomendações sobre normas de interoperabilidade pertinentes".

Dada a não previsão de contraordenações leves na esfera do altruísmo de dados (o que supomos estar associado ao facto de os dados serem disponibilizados para fins altruístas e servirem um objetivo de interesse geral, o que confere à prática do ilícito um especial juízo de censura), as contraordenações graves constantes das alíneas a) a e) do n.º 2 do artigo 8.º do DLERGD incluem, igualmente, a inobservância de formalidades que decorrem do exercício da atividade das organizações de altruísmo de dados [como a apresentação do relatório anual de atividades que, no mínimo, contenha os elementos referenciados nas alíneas a) a e)[23] do n.º 2 do artigo 20.º – cf. alínea a) do n.º 2 do artigo 8.º do DLERGD], bem como de procedimentos que garantam o nível de segurança adequado ao armazenamento e ao tratamento de dados não pessoais recolhidos com base no altruísmo de dados [cf. alínea b) do n.º 2 do artigo 8.º do DLERGD]. Também como consequência dos deveres de diligência e informação para com os titulares e/ou detentores dos dados, importa garantir uma comunicação atempada de incidentes relacionados com a transferência, acesso ou utilização não autorizados de dados não pessoais partilhados pelas organizações de altruísmo de dados [cf. alínea c) do n.º 2 do artigo 8.º do DLERGD], especificar a jurisdição do país terceiro em que a utilização dos dados se destina a ser efetuada [permitindo aos titulares dos dados dar e retirar o consentimento, e aos detentores dos dados dar e remover a autorização para o tratamento de dados, mediante a disponibilização de instrumentos próprios – alínea d) do n.º 2 do artigo 8.º do DLERGD] e respeitar as obrigações para impedir a transferência internacional ou o acesso governamental a dados não pessoais detidos na União, à margem das condições legalmente estabelecidas [alínea e) do n.º 2 do artigo 8.º do DLERGD].

Assim, ao compararmos, nos termos atrás expostos, o artigo 7.º (Contraordenações no âmbito da prestação de serviços de intermediação de dados) e o artigo 8.º (Contraordenações no âmbito do altruísmo de dados), ambos do DLERGD, percebemos uma matriz comum a partir da qual o legislador entendeu tipificar os ilícitos contraordenacionais por violação da legislação relativa à governação europeia de dados [delimitando as contraordenações muito graves a infrações a requisitos substanciais determinantes para a prestação dos serviços

[23] Os elementos essenciais a que se refere o n.º 2 são: "a) Informações sobre as atividades da organização de altruísmo de dados reconhecida; b) Uma descrição da forma como foram promovidos, durante o exercício financeiro em causa, os fins de interesse geral para os quais os dados foram recolhidos; c) Uma lista de todas as pessoas singulares e coletivas às quais foi permitido tratar os dados detidos pela entidade, incluindo uma descrição sumária dos fins de interesse geral desse tratamento e uma descrição dos meios técnicos utilizados para o realizar, nomeadamente das técnicas utilizadas para preservar a privacidade e a proteção dos dados; d) Um resumo dos resultados do tratamento de dados permitido pela organização de altruísmo de dados reconhecida, se aplicável; e) Informações sobre as fontes de receitas da organização de altruísmo de dados reconhecida, nomeadamente todas as receitas resultantes do facto de ter permitido o acesso aos dados, e sobre as despesas da entidade".

de intermediação de dados ou da atividade de altruísmo de dados, as contraordenações graves a infrações a formalidades e/ou procedimentos inerentes a deveres de informação a titulares ou detentores de dados, e obrigações de notificação e/ou reporte às autoridades competentes no âmbito das respetivas atribuições, e as contraordenações leves (apenas previstas para a prestação de serviços de intermediação de dados) para alguns factos funcionais].

No que diz respeito à responsabilidade dos infratores, o n.º 1 do artigo 9.º do DLERGD estabelece que a prática das contraordenações previstas no mencionado diploma será imputável a pessoas singulares ou a pessoas coletivas (ainda que irregularmente constituídas) e associações sem personalidade jurídica, nos termos constantes dos n.os 2 a 5 do mencionado artigo 9.º, ou seja:

1) No caso de pessoas coletivas ou entidades equiparadas, a responsabilidade pelas contraordenações pressupõe que os factos tenham sido praticados pelos titulares dos seus órgãos sociais, pelos titulares dos cargos de direção e chefia e pelos seus trabalhadores no exercício das suas funções, bem como pelas infrações cometidas pelos seus mandatários e representantes, em atos praticados em seu nome ou por sua conta[24] (cf. n.º 2 do artigo 9.º do DLERGD).

2) Caso fique provado que o agente da pessoa coletiva atuou contra ordens ou instruções expressas do ente coletivo, aquela responsabilidade ficará afastada (cf. n.º 3 do artigo 9.º do DLERGD);

3) Contudo, o legislador estabeleceu que a responsabilidade das pessoas coletivas não exclui a eventual responsabilidade individual dos agentes que possa ter lugar relativamente aos mesmos factos (cf. n.º 4 do artigo 9.º do

[24] De acordo com o teor do Parecer do Conselho Consultivo da PGR n.º 11/2013, Diário da República, 2.ª Série, n.º 178, de 16/09/2013, disponível em https://diariodarepublica.pt/dr/detalhe/parecer/11-2013-2571789, acedido em 10/09/2025: "[…] 4. O preceito do número 2 do artigo 7.º do Regime Geral das Contraordenações deve ser interpretado extensivamente, como, aliás, tem sido feito pela jurisprudência, incluindo do Tribunal Constitucional, de modo a incluir os trabalhadores, os administradores e gerentes e os mandatários ou representantes da pessoa coletiva ou equiparada, desde que atuem no exercício das suas funções ou por causa delas. 5. A responsabilidade contraordenacional das pessoas coletivas assenta numa imputação direta e autónoma, quer o fundamento dessa responsabilidade se encontre num «defeito estrutural da organização empresarial» (*defective corporate organization*) ou «culpa autónoma por défice de organização», quer pela imputação a uma pessoa singular funcionalmente ligada à pessoa coletiva, mas que não precisa de ser identificada nem individualizada". No mesmo sentido, ver o Acórdão do Tribunal da Relação de Coimbra, de 08/05/2024, Processo n.º 1079/22.0T8GRD.C1, disponível em https://www.dgsi.pt/jtrc.nsf/c3fb530030ea1c61802568d9005cd5bb/764cc3a8fc3dbd7c80258b2a0046777f?OpenDocument, acedido em 10/09/2025, cujo sumário refere: "Para a responsabilização da arguida pelo cometimento de uma contraordenação por omissão, não se exige a identificação da pessoa física que concretamente cometeu a infração, já que a responsabilidade da pessoa coletiva não é uma imputação reflexa ou indireta, mas sim autónoma".

DLERGD). Segundo Marques da Silva[25], um tema interessantíssimo "que tem vindo a ganhar relevo na doutrina é o da responsabilidade dos administradores de direito que não cumprem os seus deveres. Esta realidade não tem merecido grande atenção da doutrina e da jurisprudência, mas vem ganhando relevo porque é juridicamente inadmissível que alguém assuma uma função e voluntariamente não a cumpra. É relevante sobretudo nos casos em que da titularidade do cargo resultam deveres de agir – e são cada vez mais – e o administrador não cumpre esses deveres. Tem de ser responsabilizado pela sua omissão se se verificarem os demais pressupostos da responsabilização". Assim, contrariamente ao que sucede no n.º 3 do artigo 9.º (em que a atuação do agente individual contra ordens expressas da pessoa coletiva exclui a sua responsabilidade), esta norma admite uma corresponsabilização de pessoas singulares que exerçam funções em prestadores de serviços de intermediação de dados ou organismos de altruísmo de dados, atendendo aos especiais deveres de diligência inerentes às suas atividades.

3. Das entidades competentes para a instrução e decisão dos processos de contraordenação

Relativamente à instrução e decisão dos processos de contraordenação por violação do DLERGD, o n.º 1 do artigo 11.º estabelece que incumbe ao dirigente máximo do organismo ou das autoridades competentes designados no artigo 2.º, no âmbito das competências e poderes que lhes são legalmente atribuídos. Ainda que a Agência para a Reforma Tecnológica do Estado, I.P. (ARTE)[26] – entidade criada em 2025 no âmbito da reestruturação da Agência para a Modernização Administrativa (AMA) – não possua competências para a tramitação contraordenacional[27], nos termos da alínea b) do n.º 2 do artigo 3.º do DLERGD, cabe-lhe reportar às entidades competentes qualquer denúncia recebida ou suspeita de violação do Regulamento Governação de Dados e/ou da legislação nacional aplicável. Nos termos do n.º 2 do artigo 2.º do Decreto-Lei n.º 96/2025, de 21 de agosto, todas as referências feitas à AMA, I.P., constantes do Decreto-Lei n.º 43/2012, de 23 de fevereiro, e demais atos legislativos, regulamentares ou de outra natureza, consideram-se feitas à ARTE, I.P.

Assinala-se, ainda, que de acordo com o n.º 3 do mesmo artigo, "a colaboração entre a ARTE, I. P. e as autoridades competentes deve realizar-se nos

---

[25] Germano Marques da Silva, *Responsabilidade Penal dos Dirigentes das Sociedades*, Lisboa, Universidade Católica Editora, 2021, pp. 30-31.

[26] O Decreto-Lei n.º 96/2025, de 21 de agosto, reestrutura a Agência para a Reforma Tecnológica do Estado, I.P. (antiga Agência para a Modernização Administrativa, I.P.).

[27] As competências desta entidade constam dos n.os 1 a 3 do artigo 2.º e artigo 3.º do DLERGD.

termos do Regulamento Governação de Dados, do Código do Procedimento Administrativo e da demais legislação aplicável, promovendo-se o intercâmbio e partilha de informação, incluindo notificações e alertas, à prestação de assistência mútua, à colaboração na resolução de problemas e à aplicação mais eficaz do Regulamento Governação de Dados".

Tabela 3

Autoridades competentes para a instrução e decisão dos processos de contraordenação no Decreto-Lei n.º 2/2025, de 23 de janeiro

Autoridades competentes	
A Autoridade Nacional de Comunicações (ANACOM)	Artigo 2.º, n.º 4 do DLERGD
Entidade que sucede à Secretaria-Geral do Ministério das Finanças – ver artigo 3.º do Decreto-Lei n.º 54/2025, de 28 de março, que extingue por fusão a SGMF	Artigo 2.º, n.º 5 do DLERGD
Inspeção-Geral da Agricultura, do Mar, do Ambiente e do Ordenamento do Território	Artigo 2.º, n.º 6 do DLERGD
Autoridade de Segurança Alimentar e Económica	Artigo 2.º, n.º 7 do DLERGD
Entidade que sucede à Secretaria-Geral do Ministério do Ambiente – ver artigo 3.º do Decreto-Lei n.º 114-A/2024, de 26 de dezembro, que extingue por fusão a SGMA	Artigo 2.º, n.º 8 do DLERGD
Serviços Partilhados do Ministério da Saúde, SPMS	Artigo 2.º, n.º 9 do DLERGD
Secretaria-Geral do Governo	Artigo 2.º, n.º 10 do DLERGD

4. Da aplicação subsidiária do Regime Jurídico das Contraordenações Económicas

Em tudo o que não estiver estabelecido no Decreto-Lei n.º 2/2025, de 23 de janeiro, o artigo 13.º do DLERGD determina a aplicação subsidiária do Regime Jurídico das Contraordenações Económicas (RJCE), aprovado pelo Decreto-Lei n.º 9/2021[28], de 29 de janeiro, na sua redação atual. Esta norma materializa o disposto no n.º 2 do artigo 78.º do RJCE, ao prever que este regime se aplica subsidiariamente quando esteja expressamente determinado nos respetivos regimes jurídicos setoriais.

Verifica-se, porém, que a aplicação deste novo regime contraordenacional comum aos diferentes diplomas que regem o acesso e o exercício de atividades económicas foi concretizada através da tipificação como contraordenações económicas das situações nele abrangidas, nas alterações efetuadas pelo Decreto-Lei n.º 9/2021, de 29 de janeiro, à legislação referenciada no n.º 2 do artigo 1.º. Neste sentido, Costa e Figueiredo[29] referem que, para que se "suscite a questão da sua relação com o RJCE e, por conseguinte, a aplicação deste artigo, os regimes setoriais terão de prever um regime contraordenacional específico, isto é, tais diplomas deverão conter normas contraordenacionais económicas, seja em matéria substantiva, seja em matéria processual. […] Assim, e desde logo, os comportamentos típicos e as respetivas sanções não resultam do RJCE, mas sim dos regimes setoriais, que prevalecem sobre ele. Como tal, os factos ilícitos típicos são previstos pelos regimes setoriais e o enquadramento jurídico de qualquer infração passa, em primeiro lugar, pelos regimes setoriais e não pelo RJCE".

Acreditamos que a opção do legislador quanto a esta aplicação subsidiária do RJCE ao DLERGD decorra da manifesta desadequação do Regime Geral das Contraordenações para responder ao progresso assinalável do direito contraordenacional a que temos vindo a assistir nas últimas décadas. A este

---

[28] O Decreto-Lei n.º 9/2021, de 29 de janeiro, aprovou o Regime Jurídico das Contraordenações Económicas e promoveu alterações a uma diversidade de diplomas legais. De acordo com o preâmbulo, a sistematização efetuada no âmbito do Programa Simplex+2017, que integrou a medida Informação Contraordenacional + Simples, "permitiu avaliar a necessidade de aprovação de um regime jurídico para as contraordenações económicas, nomeadamente que estabeleça um procedimento comum e que garanta a proporcionalidade das coimas e sanções acessórias aplicáveis, que reconheça a sua especificidade e autonomia face aos demais ilícitos contraordenacionais, com o objetivo de garantir maior segurança jurídica e uniformizar e consolidar o regime contraordenacional aplicável em matéria de acesso e exercício de atividades económicas".

[29] António Augusto Costa; José Miguel Figueiredo, *Regime das Contraordenações Económicas Anotado*, Lisboa, AAFDL Editora 2021, p. 561.

propósito, Azevedo[30] considera que, dada a "evolução crescente do Direito das contraordenações em relação a cada vez mais domínios, como na economia, finanças, saúde, educação, cultura, ambiente, entre outras, e de o legislador ainda não ter estabelecido um critério distintivo entre o ilícito criminal e o ilícito contraordenacional, o Regime Geral das Contraordenações passou a ser o regime geral estagnado, sendo deixado a flutuar na ordem jurídica contraordenacional, com uma expressão diminuída". Duvidamos, porém, do carácter acertado da escolha do legislador, atendendo a que o RJCE é aplicável às contraordenações económicas[31] que sejam qualificadas por lei como tal, o que não sucedeu no Decreto-Lei n.º 2/2025, de 23 de janeiro. O que até se alcança, em virtude da falta de enquadramento de factos passíveis de preenchimento dos tipos contraordenacionais previstos no DLERGD, na noção de contraordenação económica (designadamente, no âmbito do altruísmo de dados).

5. Breves notas finais

Face à sua complexidade, a implementação do quadro legal do DLERGD configura um enorme desafio. Em primeiro lugar, pelas dificuldades na conciliação entre vários Regulamentos Europeus e a legislação nacional aplicável. De acordo com o considerando (62) do Regulamento (UE) 2022/868 (RGD), este regulamento tem como princípios orientadores o respeito pelos direitos fundamentais e a observância dos princípios reconhecidos, nomeadamente, pela Carta dos Direitos Fundamentais da União Europeia, incluindo o direito à privacidade, a proteção de dados pessoais, a liberdade de empresa, o direito de propriedade e a integração das pessoas com deficiência. No entanto, conforme referido por Novais[32], é no domínio "das relações concretas que se desenvolvem entre Administração, particulares e tribunais, que se leva a cabo a conformação prática e a concretização do conteúdo constitucional dos direitos fundamentais, pelo que, em última análise, um direito fundamental será, na vida real, aquilo que resultar do trabalho de interpretação jurídica, de aplicação, de ponderação de bens, valores e interesse conflituantes aí levado a cabo pelos diferentes operadores jurídicos, nomeadamente os juízes comuns".

---

[30] Tiago Lopes de Azevedo, *Lições de Direito das Contraordenações*, Reimpressão, Coimbra, Coimbra Editora S.A, 2022, pp. 45-46.

[31] Nos termos do n.º 2 do artigo 1.º do RJCE, "constitui contraordenação económica todo o facto ilícito e censurável que preencha um tipo legal correspondente à violação de disposições legais e regulamentares, relativas ao acesso ou ao exercício, por qualquer pessoa singular ou coletiva, de atividades económicas nos setores alimentar e não alimentar e para o qual se comine uma coima".

[32] Jorge Reis Novais, *Manual de Direitos Fundamentais*, Lisboa, AAFDL Editora, 2024, p. 568.

Em segundo lugar, pelo acréscimo de novas competências distribuídas por diferentes entidades, que terão de articular-se entre si[33], para assegurar o exercício das exigentes atribuições que lhes foram conferidas. Identificando as áreas de intervenção prioritária implementação em termos de uma governação integrada na Administração Pública, Lopes dos Santos[34] sinaliza que, ao nível organizacional, "haverá que repensar a organização segmentada e vertical existente, fomentando a criação de hierarquias horizontais de intervenção transversal. Haverá que levar em conta que a especialização de cada sector não poderá dispensar uma visão de conjunto […]. Da mesma forma, importará que a Administração planeie de forma a poder agir proactivamente na resolução de problemas complexos".

Em terceiro lugar, pela carência de recursos humanos especializados[35] aptos a darem resposta à execução das medidas exigidas. Não só pela maior rigidez nas formas de recrutamento na Administração Pública (que, na maioria dos casos, fixam como condição de candidatura a posse de relação jurídica de emprego público por tempo indeterminado), como pelas limitações salariais que decorrem da aplicação de uma tabela remuneratória específica (em função das

[33] Sobre a colaboração entre entidades, o n.º 11 do artigo 4.º do DLERGD estabelece que "as autoridades competentes e a ARTE, I. P., devem, em conformidade com as respetivas competências, cooperar de forma transversal e integrada, através do intercâmbio das informações necessárias ao exercício das suas funções no âmbito do RGD e da legislação aplicável".

[34] Mafalda Lopes dos Santos, «A Governação integrada e a Administração Pública portuguesa: 'Descompartimentar para melhor coordenar'», in Rui Marques; David Ferraz (org.), *Governação Integrada e Administração Pública*, Lisboa, INA Editora, 2015, p. 122.

[35] Nos termos do n.º 3 do artigo 7.º do RGD, "os organismos competentes devem dispor dos recursos jurídicos, financeiros, técnicos e humanos adequados para desempenhar as funções que lhes são atribuídas, inclusive dos conhecimentos técnicos necessários para poderem cumprir o direito da União ou nacional aplicável em matéria de regimes de acesso para as categorias de dados referidas no artigo 3.º, n.º 1".

respetivas carreiras e categorias), que em algumas áreas, como a das TIC[36] (Tecnologias de Informação e Comunicação), ainda não conseguem competir com o setor privado nas funções mais exigentes.

Por último, devido às desigualdades entre os estádios de desenvolvimento tecnológico das várias entidades envolvidas, que, certamente, implicarão um investimento elevado em *hardware* e *software* apropriado[37] no curto prazo. Através da Resolução do Conselho de Ministros n.º 94/2024[38], foi constituído o "Conselho para o Digital na Administração Pública". Entre outros, figuram nos objetivos deste órgão: "a) Promover o estudo da Transformação Digital na Administração Pública, incluindo a análise dos sistemas de informação e das estruturas organizacionais; b) Contribuir para a Estratégia Digital Nacional e

---

[36] Tabela remuneratória da Administração Pública na área das TIC disponível em https://www.dgaep.gov.pt/index.cfm?OBJID=b8a129f3-8eb7-4b56-932f-f084b9abab44&ID=105000000, acedido em 28/10/2025. No setor privado, de acordo com um estudo da Adecco Recruitment relativamente aos perfis mais valorizados e salários em destaque, "os perfis tecnológicos apresentam uma amplitude salarial significativa, espelhando a complexidade e o grau de especialização exigido pelo setor. No topo da tabela surge o *Cloud Engineer*, com um salário máximo de 105.000€/ano, seguido do Consultor SAP (até 98.000€) e do *Cibersecurity Engineer* (até 85.000€), refletindo a forte procura por especialistas em infraestrutura digital, segurança e soluções empresariais complexas. Funções ligadas ao desenvolvimento continuam a ser altamente valorizadas: *Mobile Developers* e *Software Engineers* podem atingir os 75.000€/ano, impulsionados pelo crescimento das plataformas digitais e pela necessidade de aplicações escaláveis e seguras. Na área da gestão e análise de dados, os *Data Engineers* têm faixas entre 32.000€ e 55.000€/ano, enquanto os *Data Analysts* oscilam entre 21.000€ e 40.000€/ano, confirmando a crescente valorização de perfis orientados para a literacia de dados e a inteligência artificial. Nos cargos de coordenação, o *IT Project Manager* mantém uma faixa estável, com salários até 70.000€/ano, e os *QA Testers* (até 60.000€) ganham destaque pela sua relevância na garantia de qualidade e na entrega de projetos tecnológicos. A base da estrutura técnica continua a ser sustentada por perfis como *Support Technician*, com salários entre 17.000€ e 35.000€/ano, e Administradores de Sistemas/Redes, que variam entre 28.000€ e 56.000€/ano, reforçando a importância de funções menos visíveis, mas críticas para a operação tecnológica das empresas. Esta diversidade salarial sublinha a especialização crescente no setor: quanto mais estratégica e técnica é a função, maior é a valorização − e maior a dificuldade em encontrar talento disponível" – disponível em https://rhmagazine.pt/adecco-recruitment-lanca-guia-salarial-para-o-setor-das-ti-para-2025/, acedido em 28/10/2025.

[37] De acordo com o Considerando (7) do RGD, existem técnicas que permitem "análises em bases de dados que contêm dados pessoais, tais como a anonimização, a privacidade diferencial, a generalização, a supressão e a aleatorização, a utilização de dados sintéticos ou similares e de outros métodos avançados de preservação da privacidade que poderão contribuir para um tratamento de dados mais favorável à privacidade". Ao que acresce o disposto no Considerando (9), quando refere que a fim de "facilitar a proteção dos dados pessoais e dos dados confidenciais e de acelerar o processo de disponibilização desses dados para reutilização ao abrigo do RGD, os Estados-Membros deverão incentivar os organismos do setor público a criarem e disponibilizarem os dados em conformidade com o princípio «abertos desde a conceção e por defeito»".

[38] Diário da República n.º 143/2024, 1.ª Série, de 25/07/2024, disponível em https://diariodarepublica.pt/dr/detalhe/resolucao-conselho-ministros/94-2024-873371754 [acedido em 28/10/2025].

o seu Plano de Ação, através da partilha dos objetivos estratégicos e necessidades setoriais específicas para a transição digital e modernização; c) Elaborar a Estratégia Digital Nacional e o seu Plano de Ação para a dimensão da digitalização dos serviços públicos e da transformação digital do Estado através da definição clara das ações, responsáveis e envolvidos, plano de implementação, indicadores de desempenho e impactos esperados". No entanto, secundando as observações de Leitão[39] "[…] Desmaterializar, digitalizar e simplificar não são, contudo, sinónimos, apesar de todas as operações serem necessárias à celeridade e eficiência da atuação administrativa. Efetivamente, a simplificação deve ocorrer previamente – e até independentemente – da desmaterialização e da digitalização, sob pena de, não sendo assim, apenas se estar a substituir a burocracia em papel por burocracia digital". Reformas estruturais com este impacto não se materializam de forma instantânea por decreto, e as mudanças necessárias exigem tempo e recursos que variam em função das prioridades orçamentais para cada ano económico, o que cria dificuldades acrescidas na execução.

[39] Alexandra Leitão, «A Administração Pública Eletrónica: oportunidades e desafios», *Cadernos de Justiça Administrativa*, n.º 152, outubro/dezembro 2022, Braga, CEJUR, pp. 24-33.

THE FRAMEWORK CONVENTION ON ARTIFICIAL INTELLIGENCE: AN AXIOLOGICAL COMPASS IN INTERPRETATION NORMS IN THE FIELD OF RIGHTS OF PEOPLE WITH DISABILITIES

*Filipe Venade de Sousa**

* Invited Adjunct Professor at the School of Education of the Polytechnic Institute of Porto and at the Lisbon School of the Faculty of Law of the Catholic University. Collaborating researcher at JusGov (Interdisciplinary Research Centre in Justice and Governance) of the School of Law of the University of Minho (Braga, Portugal). E-mail: filipesousa@ese.ipp.pt.

Abstract: An international convention specifically designed to regulate human rights in the context of Artificial Intelligence, officially called the Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law (FC-AI) was adopted. This Convention's fundamental objective is to ensure respect for human rights from the perspective of AI, in balance with technological development and innovation in various areas of daily life. This study aims to contribute to reflection on the rights of people with disabilities, in light of this Convention, which devotes special attention to them, particularly through its Article 18. This article establishes an obligation for States to take due account of the "specific needs and vulnerabilities in respect of the rights of persons with disabilities" and, therefore, refers to the provisions of the Convention on the Rights of Persons with Disabilities (CRPD), with a level of protection similar or complementary to the CRPD, according to the Framework Convention's explanatory report. The Framework Convention complements the standards established in the CRPD. Therefore, this article aims to analyze the fundamental principles and applicable standards that impact the rights of persons with disabilities. The challenges posed by artificial intelligence require a *pro homine* interpretation of the standards applied to people with disabilities, to ensure that their access to citizenship in the digital space is not hindered.

Keywords: FC-AI; CRPD; Human rights; People with disabilities; Inclusive Digital Law; *Pro homine* interpretation; Specific needs and vulnerabilities.

Resumo: Foi adotada uma convenção internacional específica para regular os direitos humanos no âmbito da Inteligência Artificial, oficialmente designado por Convenção-Quadro do Conselho da Europa sobre Inteligência Artificial, Direitos Humanos, Democracia e Estado de Direito (FC-AI). Esta Convenção tem por objeto fundamentalmente assegurar o respeito dos direitos humanos na perspetiva da inteligência artificial (IA), no equilíbrio com o desenvolvimento tecnológico e inovação nos diferentes domínios da vida quotidiana. O presente artigo visa contribuir para a reflexão sobre os direitos das pessoas com deficiência, à luz desta Convenção, que lhes dedica especial atenção, nomeadamente através do seu artigo 18.º. Este artigo estabelece uma obrigação para os Estados que tenham em conta devidamente as "necessidades e vulnerabilidades específicas relativamente ao respeito dos direitos das pessoas com deficiência"

e, por isso, remete às disposições da Convenção sobre os Direitos das Pessoas com Deficiência (CDPD), com um nível de proteção similar ou complementar à CDPD, segundo informa o relatório explicativo da Convenção-Quadro. A Convenção-Quadro é complementar com as normas estabelecidas na CDPD. Assim, o presente artigo propõe-se analisar os princípios fundamentais e as normas aplicáveis que impactem os direitos das pessoas com deficiência. Os desafios lançados pela inteligência artificial não deixam de exigir uma interpretação *pro homine* das normas aplicadas às pessoas com deficiência, para lhes assegurar que não lhes seja obstaculizado o acesso à cidadania no espaço digital.

Palavras-chave: FC-AI; CDPD; Direitos humanos; Pessoas com deficiência; Direito Digital Inclusivo; Interpretação *pro homine*; Necessidades e vulnerabilidades específicas.

"For many citizens, scientific discoveries have alleviated the suffering caused by disease or disability. Nonetheless, the prospect of gaining such valuable scientific knowledge need not and should not be pursued at the expense of human rights or human dignity."[1]

[1] *Cf.* R. Brownsword, *Rights, regulation, and the technological revolution*, Oxford University Press, 2008, p. 33.

1. Introduction

History reveals that every revolution — social, political, cultural, and currently, technological — rapidly transforms the domains of daily life.[2] However, technological revolutions, including Artificial Intelligence (AI) systems, do not always keep pace with the legitimate expectations of people with disabilities. Frequently, their needs and rights are not among the main concerns when discussing technological advancements. The rapid development of new technologies, particularly AI systems, therefore raises pressing questions about the need for new legal responses that guarantee the full inclusion of people with disabilities.[3]

As Quintavalla & Temperman rightly state, AI systems can be an "emancipatory tool"[4] for people with disabilities. Technological advances simultaneously offer opportunities and risks to facilitate and improve accessibility in general, while also addressing risks related to the exercise of fundamental rights and freedoms. For example, "In fact, many of the AI-based technologies that are either specifically or tangentially developed to assist PWDs [people with disabilities] are built according to predefined and rigid logic rules that do not capture the nuances of the notion of disability and, in addition, they might be built based on ableist assumptions."[5]

The fundamental question is whether technological advances change rapidly, whether people with disabilities accept that technological solutions, including AI systems, are obsolete or outdated and do not meet their legitimate expectations, and how does the law keep up with technological advances that evolve rapidly in each social life cycle?[6] As Askland warns, "If legislation and regulation seek to provide legal guidance before the fact, and adjudication seeks to classify hardships after the fact, then legal actions will surely be misplaced unless they are adequately informed by the relevant facts."[7] However, according to the same author, "As research nears application, the law should take note, but only as it anticipates the extension of familiar jurisdictional definitions to new

---

[2] Luciano Floridi, *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*, Oxford University Press UK, 2014, p. 1; Liane Colonna & Stanley Greenstein (eds.), *Law in the Era of Artificial Intelligence*, Nordic Yearbook of Law and Informatics 2020–2021, 2022.

[3] Antonella Zarra; Silvia Favalli &Matilde Ceron, «Artificial Intelligence and Disability Rights», in Alberto Quintavalla & Jeroen Temperman, *Artificial Intelligence and Human Rights*, Oxford University Press, 2023, p. 248.

[4] *Ibidem*, p. 251.

[5] *Ibidem*, p. 250.

[6] Andrew Askland, «Introduction: Why Law and Ethics Need to Keep Pace with Emerging Technologies», in Gary Marchant et al. (eds.), *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight: The Pacing Problem*, Spring 2011, p. xiii.

[7] *Ibidem*, p. xv.

subject matter. Law needs not worry about its ignorance of the subject matter or the methodologies of the research in its formative stages save for focused concerns, e.g., about informed consent for human subjects or intrinsically dangerous subject matter (and these concerns do re-quire considerable attention)."[8]

This study analyzes the legal frameworks related to the rights of people with disabilities in light of the Council of Europe Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law (FC-AI) and the United Nations Convention on the Rights of People with Disabilities (CRPD), as well as the interpretation of the FC-AI through the explanatory report that complements the understanding of this study. AI Law expands the new legal boundaries of Technology Law and Disability Law, and indeed, we are still in "uncharted territory" when it comes to clarifying whether the rights of people with disabilities can be guaranteed under conditions of equal treatment, not just in legal terms, but also in social, doctrinal, and jurisprudential terms. We explain below.

[8] *Ibidem*, p. xvi; Wolfgang Hoffmann-Riem, «Artificial Intelligence as a Challenge for Law and Regulation», in Thomas Wischmeyer & Timo Rademacher (eds.), *Regulating Artificial Intelligence*, Springer International Publishing, 2020, pp. 1–29; Ubena John, «Can Artificial Intelligence be Regulated? Lessons from Legislative Techniques», in *Nordic Yearbook of Law and Informatics 2020–2021 Law in the Era of Artificial Intelligence*, The Swedish Law and Informatics Research Institute, 2022, pp. 273–94.

2. Inclusive Digital Law: Correlation between AI, Technology Law, and Disability Law

The legal regulation of AI in the context of the rights of people with disabilities requires a contribution to the construction of *Inclusive Digital Law* (IDL)[9] that responds to the new legal and political boundaries of today in the context of new technologies that challenge traditional legal understandings.[10]

The main question of IDL is how to ensure that technological advances do not exacerbate existing inequalities for people with disabilities and how to promote digital and technological inclusion — not only in the social and human dimensions — and, above all, whether they ensure the right to individual freedom and personal autonomy in the context of technological advances.[11] In this sense, the dogmatic construction of IDL that responds to these current challenges is necessarily anchored in the multi-level legal frameworks of the CRPD and the FC-AI.

We define IDL as the fundamental right of every person, and particularly people with disabilities, to access, participate, and self-determine in any digital and technological environment — including those mediated by AI — under conditions of equality and dignity, through guarantees of accessibility, personal autonomy, and individual freedom, imposing positive obligations on the State and private actors regarding inclusive design, technological adaptation, and protection against exclusions and undue conditioning.

---

[9] The concept of IDL is inspired and adapted from the report published by EUD, "Sign Language in the Era of Artificial Intelligence" (2025, pp. 51–52), which defines Deaf Digital Law: "The Deaf Digital Law idea introduces a novel legal framework for the full integration of deaf people into cyberspace. It broadens the common understanding of digital accessibility to cover, for instance, the implementation of sign languages in digital interfaces, virtual assistants, sign language inclusion into social networking sites, streaming services, and customer service platforms."

[10] The studies inspired by the present study were: Silja Voeneky; Philipp Kellmeyer; Oliver Mueller & Wolfram Burgard, *The Cambridge Handbook of responsible artificial intelligence. Interdisciplinary perspectives*, Cambridge University Press & Assessment, 2022; Nathalie A. Smuha, *The Cambridge Handbook of the Law, Ethics and Policy of Artificial Intelligence*, Cambridge University Press, 2025; Larry A. Dimatteo; Cristina Poncibò & Michel Cannarsa, *The Cambridge Handbook of Artificial Intelligence. Global perspectives on law and ethics*, Cambridge University Press, 2022; R. Valle Escolano, «Artificial intelligence and rights of people with disabilities: The power of algorithms», *Revista Española de Discapacidad*, 11(1), 2023, pp. 29–49, https://doi.org/10.5569/2340-5104.11.01.03; M. Balcerzak & J. Kapelańska-Pręgowska (eds.), *Artificial Intelligence and International Human Rights Law*, Cheltenham, UK: Edward Elgar Publishing, Retrieved Sep. 12, 2025, from https://doi.org/10.4337/9781035337934.

[11] C. El Morr; B. Kundi; F. Mobeen; S. Taleghani; Y. El-Lahib & R. Gorman, «AI and disability: A systematic scoping review», *Health Informatics Journal*, 30(3), 2024, doi:10.1177/14604582241285743; Filippo A. Raso; Hannah Hilligoss; Vivek Krishnamurthy; Christopher Bavitz & Levin Kim, *Artificial Intelligence & Human Rights: Opportunities & Risks*, Research Publication No. 2018-6.

Within the scope of the IDL, we can speak of the type of AI Law that specializes the legal frameworks of Technology Law.[12] AI Law, like any Technology Law[13], and especially the IDL, encompasses, on the one hand, human dignity and personal autonomy — that is, the right to regulate this sphere of daily life, free from digital and social interference — and, on the other, the right not to be conditioned by services provided through digital or technological platforms. Thus, the right to self-determination in the inclusive digital sphere is characterized as the right of people with disabilities to define and select what is best for their legitimate interests within the framework of the digital and technological solutions available. In this context, the right to self-determination is assumed as a right to freedom, the freedom to inform or communicate, without conditions, one's intentions within the inclusive digital sphere. Without this element, individual freedom would be curtailed in the freedom to be able to choose whoever you want, whenever you want, for as long as you want and as many times as you want.

Thus, the proposed definition of IDL can be considered a specialized and complementary area of Technology and Disability Laws, seen as a multidimensional legal platform supported by three fundamental pillars. Thus, the three-dimensional structure of IDL is characterized by the following: first, IDL consists of affirming technological (or digital) *accessibility* as a *right-means*. AI systems must be compatible with the different specific needs of people with disabilities and must also be adaptable, in any context, to the technical or functional specificities of people with disabilities who can access them, without any discriminatory obstacles. Second, *personal (digital) autonomy* as a *right-faculty*. Personal autonomy implies that each and every people with disability has the power — the decision-making power over the means and forms to choose — over when, how, and under what conditions to interact with AI systems. It includes the possibility of rejecting technological solutions incompatible with the legitimate needs or preferences of people with disabilities. Third, *individual (digital) freedom* as a *right-end*. Individual freedom implies that people with disabilities have the full enjoyment and exercise of their rights to be, to exist, to design, to choose, and to participate, within the framework of personal self-determination, in the digital space, without undue conditioning or arbitrary restrictions. This means that there can be no algorithmic conditioning or arbitrary dependence due to the choices of AI systems that do not have the consent or will of the people with disability. These three interdependent elements contribute to the effectiveness of

[12] European Platform for Rehabilitation (EPR), *EPR Report on "AI and New Technologies for Inclusion Annual Conference"*, 2024, available at: https://www.edf-feph.org/blog/ai-and-new-technologies-in-services-for-persons-with-disabilities/ [accessed 9/5/2025].

[13] María Emilia Casas Baamonde & Daniel Pérez del Prado, *Derecho y Tecnologías*, Editorial Centro de Estudios Ramón Areces, S.A., 2025.

the technological and digital inclusion of people with disabilities, who exercise their rights with equal treatment with others.

3. The FC-AI as an essential parameter for the rights of people with disabilities in the area of AI

3.1. *Purpose of the FC-AI*

In 2024, for the first time, within the Council of Europe, the various States adopted the *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law* (FC-AI).[14] There is relevant legal literature on the impact of FC-AI.[15]

The final text of the FC-AI was adopted on May 17, 2024. Unprecedentedly, the new Convention can be seen as a Convention with a trans-European vocation, to the detriment of being specifically European, as it is open for signature by any non-member State that participated in the process of negotiation and adoption of the FC-AI, in addition to the member States of the Council of Europe, for example, the European Union itself as a signatory to the Convention, the United States, Canada, among other countries interested in signing the new Convention, in accordance with Articles 30 and 31 of the FC-AI.

---

[14] *The Framework Convention on Artificial Intelligence*, available at: https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence [accessed 9/5/2025].

[15] M. Rotenberg, «Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (Council Eur.)», *International Legal Materials*, 64(3), 2025, pp. 859–902, doi:10.1017/ilm.2025.1; Cheng-chi (Kirin) Chang, «The First Global AI Treaty: Analyzing the Framework Convention on Artificial Intelligence and the EU AI Act», *2024 University of Illinois Law Review* (Online), 86, 2024, available at: https://ssrn.com/abstract=5069335 [accessed 9/5/2025]; José Jaime Villalobos & Matthijs M. Maas, «Beyond a Piecemeal Approach: Prospects for a Framework Convention on AI' (24 June 2025)», in Philipp Hacker et al. (eds), *The Oxford Handbook of the Foundations and Regulation of Generative AI*, online edn, Oxford Academic, 22 Apr. 2025, available at: https://doi.org/10.1093/oxfordhb/9780198940272.013.0044 [accessed 9/5/2025]; Philipp Hacker et al. (eds.), *The Oxford Handbook of the Foundations and Regulation of Generative AI*, online edn, Oxford Academic, 22 Apr. 2025, available at: https://doi.org/10.1093/oxfordhb/9780198940272.001.0001 [accessed 9/5/2025]; Duncan Cass-Beggs, Stephen Clare, Dawn Dimowo & Zaheed Kara, «Framework Convention on Global AI Challenges», CIGI Discussion Paper June 25, 2024; Karolína Babická & Cristina Giacomin, «Understanding the Scope of the Council of Europe Framework Convention on AI», available at: https://opiniojuris.org/2024/11/05/understanding-the-scope-of-the-council-of-europe-framework-convention-on-ai/ [accessed 9/5/2025]; M. Á. Presno Linera & A. Meuwese, «Regulating AI from Europe: a joint analysis of the AI Act and the Framework Convention on AI», in *The Theory and Practice of Legislation*, 2025, pp. 1–20, available at: https://doi.org/10.1080/20508840.2025.2492524 [accessed 9/5/2025]; Sabrina Ghielmini; Christine Kaufmann; Charlotte Post; Tina Büchler; Mara Wehrli & Michèle Amacker, *Grund- und Menschenrechte in einer digitalen Welt / Droits fondamentaux et droits humains à l'ère numérique*, Bern: tredition/buch netz, 2021.

The purpose and objective of the FC-AI are to ensure that AI activities are fully compatible with human rights, democracy, and the rule of law.[16] The Convention itself does not regulate all aspects of AI activities, nor issues related to AI technologies as such. In other words, the purpose and objective of the Convention is specifically targeted at the Council of Europe's mandate, with a particular focus on the potential interference of AI systems with human rights, democracy, and the rule of law.[17] As a result, the FC-AI establishes regulating state obligations in the field of AI and does not create new rights established in that Convention nor reduce the scope and content of human rights provided for by other legal instruments in the field of human rights.[18]

3.2. *Scopes of application of the FC-AI*

The normative structure of the Convention is divided into different chapters throughout the Convention. The definition of the general obligations of the "public authorities" of the respective States can be found in the general framework provided for in Chapter II of the Convention.

Article 4 defines two essential terms: "adoption" and "maintenance" of measures aimed at ensuring that activities carried out in the field of AI "are consistent with obligations to protect human rights, as enshrined in applicable international law and in its domestic law," within the framework of multilevel multinormativity.[19] Consequently, States have the obligation to ensure that their legislative measures comply with their applicable international obligations and, likewise, they have the obligation to adopt, within the scope of political and legislative conformity, any necessary measures to ensure respect for human rights in the field of AI.[20]

The concept of state obligations established by the Convention implies that it defines the various state obligations that ensure a "comprehensive approach towards addressing AI-related risks and adverse impacts on human rights, democracy and the rule of law by capturing all stages of activities relevant to artificial intelligence systems."[21] Furthermore, the concept of obligations also provides for the determination of obligations that ensure the prevention of

---

[16] *The Framework Convention on Artificial Intelligence*, Explanatory Report – CETS 225 – Artificial Intelligence, 5.IX.2024, § 12.

[17] *Ibidem*, § 12.

[18] *Ibidem*, § 13.

[19] *Ibidem*, §§ 16-17.

[20] *Ibidem*, §§ 37-41.

[21] *Ibidem*, § 14.

future, not just current, risks. This means that the FC-AI is an evolving and dynamic legal instrument that keeps pace with new technological realities.

According to Article 3, the scope of the Convention has a general regime and also a regime of exceptions. Starting with the latter, the exceptions to the application of the Convention cover three areas not applicable to the Convention:

> *(i) Article 3(2): "protection of their national security interests, provided that such activities are carried out in a manner consistent with applicable international law, including obligations arising from international human rights law, and with respect for their democratic institutions and processes.";*
>
> *(ii) Article 3(3): "research and development activities related to artificial intelligence systems that have not yet been made available for use, unless testing or similar activities are carried out in a manner that is likely to undermine human rights, democracy and the rule of law.";*
>
> *(iii) Article 3(4): "Matters relating to national defense".*

The general application regime covers "activities within the life cycle of artificial intelligence systems that may undermine human rights, democracy, and the rule of law" (Article 3(1)) carried out by "public authorities," (Article 3(1)(a)) which are direct recipients of the Convention's application, or "by private actors acting on their behalf," (Article 3(1)(a)) which are indirect recipients of the Convention's application, carrying out these activities on behalf of "public authorities" responsible for the acts and omissions of "private actors" that violate the Convention's provisions.[22] To fulfill their obligations under the Convention, "public authorities" must address the "risks and impacts arising from activities carried out in the life cycle of artificial intelligence systems by private actors" (Article 3(1)(b))[23] in accordance with the Convention's objective and purpose.

3.3. *FC-AI and the rights of people with disabilities*

3.3.1. General considerations

The legal treatment of AI has already been extensively addressed by soft law instruments, at varying levels[24], including declarations and recommenda-

[22] *Ibidem*, § 28.

[23] *Ibidem*, § 29.

[24] Clémentine Pouzet, *Les modes de régulation de l'intelligence artificielle par le droit européen*, Droit, 2020, available at: https://dumas.ccsd.cnrs.fr/dumas-04261774v1 [accessed 9/5/2025].

tions from, among others, the Council of Europe[25], the OECD[26], and UNESCO[27], and, more recently, at the European Union level, European regulations on AI.[28]

In general, the UN recognizes that AI represents a significant technological advance for people with disabilities. The report considers that AI-based technologies "appear to be a mix of promise, progress, and challenges."[29] A report was presented analyzing the impacts and risks of AI for people with disabilities, particularly in the design of AI algorithms, the lack of inclusion and accessibility of AI, and the discrimination against people with disabilities by AI.

3.3.2. Purpose and scope of Article 18

The FC-AI contains a specific rule regarding the rights of people with disabilities in Article 18, without prejudice to other interconnected articles throughout the conventional text. This article establishes the following:

---

[25] For example, Recommendation of the Committee of Ministers of the Council of Europe to Member States on the human rights impacts of algorithmic systems, adopted on 8 April 2020 - CM/Rec(2020); *Les droits humains dès la conception de l'IA – une protection durable des droits humains à l'ère de l'intelligence artificielle*, 2023.

[26] Recommendation on Artificial Intelligence adopted by the OECD Council on 22 May 2019 (the "OECD AI Principles").

[27] UNESCO Recommendation on the Ethics of Artificial Intelligence adopted on 23 November 2021.

[28] EU Regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) adopted on 13 March 2024; Lorenzo Cotino Hueso & Pere Simón Castellano (dirs.), *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea*, Editorial Aranzadi, S.A.U., 2024.

[29] Conférence des États Parties à la Convention relative aux droits des personnes handicapées, *Ne laisser personne de côté : utiliser l'intelligence artificielle comme moyen de favoriser l'inclusion pour renforcer la participation des personnes handicapées*, 2025, Doc. CRPD/CSP/2025/3; Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, David Kaye, on *AI technologies and implications for the information environment* (2018), Doc. A/73/348; Human Rights Council, *Rights of persons with disabilities*, Report of the Special Rapporteur on the rights of persons with disabilities, A/HRC/49/52, 2022; Virginia Carcedo & Yllera Jesús Hernández Galán, «Estudio del impacto de la Inteligencia Artificial en los derechos de las personas con discapacidad», Resolución aprobada por la Asamblea General el 22 de septiembre de 2024, A/RES/79/1.

Article 18 – Rights of persons with disabilities and of children

Each Party shall, in accordance with its domestic law and applicable international obligations, take due account of any specific needs and vulnerabilities in relation to respect of the rights of persons with disabilities and of children.
The origins of the formulation of Article 18 of the FC-AI have generally been little discussed.[30] *There are some contributions to the FC-AI project that seek to integrate specific considerations on the rights of people with disabilities, consolidating the origin and conventional formulation of the specific standard for these peoples. The same article can be understood by the "explanatory report" in the following way. The interpretation of the Convention through the explanatory report does not constitute an "instrument providing an authoritative interpretation of the text of the Framework Convention although it may facilitate the understanding of its provisions."*[31]

Article 18 – Rights of persons with disabilities and of children

116. This provision sets out an obligation for the Parties, in the context of the activities within the lifecycle of artificial intelligence systems, to take due account of "specific needs and vulnerabilities in relation to respect of the rights of persons with disabilities and of children" and in this connection it contemplates directly to the provisions and the legal regime of the UNCRPD and the UNCRC as well as the applicable domestic law of each Party on the rights of persons with disabilities and the rights of the child. Explicit reference to the applicable domestic law on the rights of the child and the rights of persons with disabilities has been inserted, in particular, with a view to take into consideration the situation of any Party to the Framework Convention which did not ratify the UNCRC or the UNCRPD, but nevertheless has the domestic legislation securing the enjoyment of such rights.
117. The reference to the domestic law in this provision is meant solely to point at provisions of domestic law which provide the level of protection in the relevant context similar or supplementary to the UNCRPD or the UNCRC, and such reference cannot be invoked by a Party as justification for its failure to perform this treaty obligation. The objective is thus to guarantee the highest possible level of consideration for any specific needs and vulnerabilities in relation to respect of the rights of persons with disabilities and of children, including training on digital literacy, as explained in relation to Article 20 in the Explanatory Report.
118. In view of the serious risk that artificial intelligence technologies could be used to facilitate sexual exploitation and sexual abuse of children, and the specific

[30] Council of Europe, Committee of Ministers, «Draft framework convention on artificial intelligence, human rights, democracy and the rule of law», Doc. 15951, 2024, available at: https://pace.coe.int/en/files/33424/html [accessed 9/5/2025].

[31] *The Framework Convention on Artificial Intelligence*, *op. cit.*, § II.

> *risks that it poses to children, in the context of implementation of this provision, the Drafters considered the obligations set forth in the Lanzarote Convention, the Optional Protocol to the UN Convention on the Rights of the Child on the sale of children, child prostitution and child pornography, and General Comment No. 25 to the UNCRC on children's rights in relation to the digital environment.*

3.3.3. The high level of protection of the rights of people with disabilities: Correspondence between the FC-AI and the CRPD

Articles 21 and 22 of the FC-AI and Article 4(4) of the CRPD share the same perspective of *pro homine* application for the sake of enhanced protection of the rights of people with disabilities.[32] The meaning of the "wider measure of protection" is equated with the principle of *pro homine* interpretation. That is, it "can be interpreted as providing the possibility of putting a person, for example, in a more favorable position than provided for under the Framework Convention."[33]

This is a complementarity in the protection of rights between the CRPD and the FC-AI. As a result, these articles target an essential and common parameter for the protection of rights. It is true that these provisions complement each other for the purposes of interpreting standards relating to the rights of people with disabilities. However, if (inter)national legislation reveals its rules for protecting protected rights, it may prevail. The explanatory report considers that the application of national legislation is only legitimate if it has "the level of protection in the relevant context similar or supplementary to the UNCRPD or the UNCRC, and such reference cannot be invoked by a Party as justification for its failure to perform this treaty obligation."[34] Consequently, the level of similar or complementary protection is already recognized within international human rights law and is described in Articles 21 and 22 of the FC-AI, meaning that the provisions of the FC-AI "harmoniously coexists with other international human rights treaties and instruments."[35]

The implementation of the FC-AI allows for "certain margin of flexibility as to how exactly to give effect to the provisions of the Framework Convention, in view of the underlying diversity of legal systems, traditions and practices

---

[32] Filipe Venade de Sousa, «The Charter of Fundamental Rights of the European Union and the United Nations Convention on the Rights of Persons with Disabilities: a dynamic pro unione and pro homine with particular reference to the CJEU case-law», *UNIO – EU Law Journal*, *5*(1), 2019, pp. 109–120, available at: https://doi.org/10.21814/unio.5.1.255 [accessed 9/5/2025].

[33] *The Framework Convention on Artificial Intelligence*, *op. cit.*, § 127.

[34] *Ibidem*, § 117.

[35] *Ibidem*, § 125.

among the Parties and the extremely wide variety of contexts of use of artificial intelligence systems in both public and private sectors."[36] Throughout the Convention text, the expression "adopt or maintain" is frequently found. This implies that States have a margin of flexibility to fulfill their obligations by "adopting new measures or by applying existing measures such as legislation and mechanisms that existed prior to the entry into force of the Framework Convention. Use of both of these terms acknowledges that, for the purpose of domestic implementation, either of these approaches may be equally sufficient."[37]

The adoption or maintenance of "graduated and differentiated" measures establishes that any measures adopted must be "tailored to the level of risk posed" by an AI system, "within specific spheres, activities and contexts, as appropriate, and that this task falls on Parties to the Framework Convention to decide how to balance the relevant competing interests in each sphere, taking into account specificities of activities in the private sector, their domestic regulatory framework and national agenda for artificial intelligence while ensuring the protection and promotion of human rights, democracy and the rule of law. The Parties may also take into account specificities of public sector activities such as law enforcement, migration, border control, asylum and the judiciary."[38]

3.3.4. The scope of AI and new technologies

Article 1 of the FC-AI aims to ensure that all activities related to the use of AI must be "fully consistent" with human rights. Therefore, government authorities must adopt all necessary measures, whether legislative, policy, or administrative, to fulfill their obligations under the Convention. These measures must be "graduated and differentiated as necessary"[39] based on the type and degree of severity and likelihood of adverse impacts on human rights resulting from the use of AI. The compliance of measures related to the control of risks of AI use that have the potential to interfere with human rights is provided for in Article 3 of

[36] *Ibidem*, § 16.

[37] *Ibidem*, § 17.

[38] *Ibidem*.

[39] *Ibidem*.

the Convention, and Article 4 further stipulates that the measures adopted must be in accordance with the obligations under international human rights law.[40]

The concept of AI is defined in Article 2 of the FC-AI. The definition of "artificial intelligence system" in Article 2 of the FC-AI describes AI as a technical concept: machine-based systems in general. It is a general, evolutionary and flexible concept.[41] This definition must also be read in light of the CRPD, which has several ramifications in the use of AI with the "potential to interfere with human rights," and the concept of AI takes into account "the graduated and differentiated approach" and "contextual elements in the Framework Convention's individual provisions"[42] in conjunction with the provisions of the CRPD that regulate the rights of persons with disabilities in relation to AI.

This definition provides a minimum conceptual framework that is articulated with the concepts of "universal design," "accessibility," and "reasonable accommodations" within the meaning of Articles 2 and 9 of the CRPD. These articles offer ethical and legal criteria for how this technology should be designed: accessible, inclusive, and promoting equality. In this sense, AI systems cannot simply be efficient or innovative; they must also respect human diversity. Universal design, applied to AI, reminds us that autonomy and technological adaptability should serve to integrate, not to exclude.

3.3.5. Fundamental principles in the dialogical normative framework between the FC-AI and the CRPD

Article 6 of the FC-AI establishes the "common general principles" that States must adopt a "general approach" to implementing measures related to activities carried out in the life cycle of AI systems, "in line with the approach described in Article 1, paragraph 2, in a manner appropriate to its domestic le-

---

[40] There are several published studies on the use of AI within the Council of Europe: Florence G'sell & Florian Martin-Bariteau, «The impact of Blockchains for Human Rights, Democracy, and the Rule of Law An Overview», Report to the Council of Europe, 2022; David Leslie; Christopher Burr; Mhairi Aitken; Josh Cowls; Mike Katel & Morgan Briggs, «Feasibility study on a legal framework on AI design, development and application based on Council of Europe's standards», adopted by the CAHAI on 17 December 2020; CAHAI – Ad hoc Committee on Artificial Intelligence, «Towards regulation of AI Systems Global perspectives on the development of a legal framework on Artificial Intelligence (AI) systems based on the Council of Europe's standards on human rights, democracy and the rule of law», Council of Europe, December 2020; Karen Yeung «A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework», DGI(2019)05.

[41] *The Framework Convention on Artificial Intelligence*, *op. cit.*, §§ 23-24.

[42] *Ibidem*, § 25.

gal system, and also taking into account the other obligations contained in this Framework Convention."[43] Indeed, the principles contained in the Convention, which are drafted with a "high level of generality, with the intention that they should be overarching requirements that can be applied flexibly in a variety of rapidly changing contexts. They are also purposive, expressing the reason behind the rule and have very broad application to a diverse range of circumstances."[44]

The FC-AI takes into account the principles and standards set forth by the CRPD, as stated in the FC-AI preamble. The preamble to the same Convention states that the FC-AI: "Recognising the framework character of this Convention, which may be supplemented by further instruments to address specific issues relating to the activities within the lifecycle of artificial intelligence systems;" and "Mindful also of the 1989 United Nations Convention on the Rights of the Child and the 2006 United Nations Convention on the Rights of Persons with Disabilities".

[43] *Ibidem*, § 50.

[44] *Ibidem*, § 49.

Table 1: Fundamental principles regarding AI

FC-AI	CRPD
Article 7 – Human dignity and individual autonomy Article 8 – Transparency and oversight Article 9 – Accountability and responsibility Article 10 – Equality and non-discrimination Article 11 – Privacy and personal data protection Article 12 – Reliability Article 13 – Safe innovation	Article 3 - General principles The principles of the present Convention shall be: (i) Respect for inherent dignity, individual autonomy including the freedom to make one's own choices, and independence of persons; (ii) Non-discrimination; (iii) Full and effective participation and inclusion in society; (iv) Respect for difference and acceptance of persons with disabilities as part of human diversity and humanity; (v) Equality of opportunity; (vi) Accessibility; (vii) Equality between men and women; (viii) Respect for the evolving capacities of children with disabilities and respect for the right of children with disabilities to preserve their identities.

Note: Own elaboration

3.3.5.1. The specific principles of the FC-AI: technical and ethical aspects

The specific principles related to the use of AI must be incorporated into the Parties' domestic approaches to the regulation of AI systems. These principles are also based on their respective international human rights obligations, which have a "detailed legal regime of human rights protection with its own set of rules, principles and practices regarding the scope, content of rights and possible restrictions, derogations or exceptions to these rights as well as the functioning of the applicable supervision and enforcement mechanisms."[45] As a result, the application of the FC-AI principles must be aligned with the principles and standards established by the CRPD regarding the rights of people with disabilities. The explanatory report clearly emphasizes that "Furthermore, nothing in this Framework Convention is intended to impact existing human rights obligations whenever they overlap with the principles in Chapter III."[46]

The specific principles of the FC-AI relating to activities carried out within the scope of AI cover, in particular, the principles of transparency and oversight under Article 8. On the one hand, it is required that the "transparency and oversight requirements" be "appropriate to specific contexts and risks" and that the activities carried out (including any activities, identification of content generated by AI systems) "are in place" that allows for the analysis of whether they comply with these essential requirements. Furthermore, the principles of responsibility and accountability described in Article 9 require the existence of "accountability for adverse impacts on human rights, democracy and the rule of law" resulting from activities carried out within the scope of AI. Respect for privacy and the protection of personal data implies that States adopt or maintain necessary measures, of various legal natures, in accordance with applicable national and international standards, to guarantee the protection of "privacy rights" regarding personal data, as well as effective safeguards for their protection, as per Article 11 of the FC-AI.

Another relevant principle is trustworthiness (Article 12). States take "appropriate measures" to promote the trustworthiness of AI systems that benefit from "confidence in their results," and must meet, at a minimum, "adequate quality and safety requirements" throughout the lifecycle of AI systems.

Safe innovation is the relevant principle of the FC-AI, which states that innovations related to systems developed by AI must be safe to "avoid negative impacts on human rights, democracy, and the rule of law," according to Article 13. This principle allows States to establish, "where appropriate," "controlled en-

[45] *Ibidem*, § 51.

[46] *Ibidem*, § 52.

vironments for the development, experimentation, and testing of artificial intelligence systems under the supervision of their competent authorities."

For human rights violations resulting from activities carried out within the scope of AI, there must be "accessible and effective remedies" to redress any violations, pursuant to Article 14 of the FC-AI. Furthermore, there must be procedural guarantees to ensure "effective safeguards, protections, and due process rights," in accordance with applicable international and national law, according to Article 15 of the same Convention, so that any activities carried out by AI have a "significant impact" on the enjoyment of the human rights of "affected persons." Likewise, States "shall seek to ensure" that, "as appropriate to the context," people interacting with AI systems must be informed that "they are interacting with such systems and not with a human being."

The Convention requires States to carry out risk and negative impact assessments and mitigation procedures under Article 16. This risk assessment takes into account the principles established by the Convention itself. Among other obligations, States essentially adopt measures to identify, assess, prevent, and mitigate the risks posed by AI systems, "taking into account the actual and potential impacts" on human rights, democracy, and the rule of law.

3.3.5.2. Axiological principles of the use of AI in light of the CRPD: framing the protection of dignity, equality and non-discrimination

The principles established by the FC-AI take into account the particularities of the rights of people with disabilities, which are respected and guaranteed in all contexts. In any case, AI systems or technologies must be appropriate and proportionate to achieve the legitimate purposes pursued for the benefit of an inclusive society and, therefore, must not violate the values and principles prescribed by the CRPD.

(a) Human dignity and personal autonomy

Article 7 of the FC-AI emphasizes the importance of human dignity and individual autonomy: "The Drafters also wanted to draw attention to human dignity and individual autonomy as foundational values and principles that are essential for the full realisation of human rights, democracy and the rule of law and that can also be adversely impacted by certain activities within the lifecycle of artificial intelligence systems."[47]

[47] *Ibidem*, § 10.

The principles of human dignity and individual autonomy are "part of human-centric regulation and governance of the activities within the lifecycle of artificial intelligence systems that fall in the scope of the Framework Convention."[48]

This article establishes that States shall "adopt" or "maintain measures" that respect these principles. The inherent dignity of the human person is the underlying basis of human rights. The dignity of each and every people with disabilities is respected and inviolable, and every human being has the right to respect for human dignity.[49] Each and every people cannot be "objectified" and is merely an end in itself.[50] No people with disability may be harmed, in any context, by AI systems, and there may be no violation of inherent dignity or abuse of fundamental rights and freedoms. Taking these principles into account, AI-driven activities "should not lead to the dehumanization of individuals, undermine their agency or reduce them to mere data points, or anthropomorphise artificial intelligence systems in a way which interferes with human dignity. Human dignity requires acknowledging the complexity and richness of human identity, experience, values, and emotions."[51]

On the other hand, the principle of individual autonomy is "one important aspect of human dignity and refers to the capacity of individuals for self-determination; that is, their ability to make choices and decisions, including without coercion, and live their lives freely."[52] In other words, individual autonomy implies that people with disabilities themselves have the right to define, choose, and control, within the framework of their personal self-determination, the use and impact of AI-driven technologies. Therefore, the right to define and control these activities cannot be undermined in favor of AI. For better understanding, "Human-centric regulation acknowledges the significance of allowing individuals to shape their experiences with artificial intelligence, ensuring that these

---

[48] *Ibidem*, § 53.

[49] Filipe Venade de Sousa, «A multifuncionalidade da dignidade da pessoa humana e as pessoas com deficiência», *Scientia Iuridica – Revista de Direito Comparado Português e Brasileiro*, n.º 349, 2019; Filipe Venade de Sousa, «Acquis conventionnel da dignidade inerente às pessoas com deficiência», in A. Sofia Pinto Oliveira & Patrícia Jerónimo (eds.), L*iber Amicorum Benedita MacCrorie, Volume I*, Portugal: UMinho Editora, 2023; Filipe Venade de Sousa, *Direitos Fundamentais das Pessoas com Deficiência e Jurisprudência Multinível*, Universidade Católica Editora, Lisboa, 2021.

[50] Filipe Venade de Sousa, «Direito das Pessoas com Deficiência», in J. E. Franco; P. Jerónimo; S. M. Alves-Jesus & T. C. Moreira (coords.), *Dicionário Global dos Direitos Humanos*, 2022, available at: https://dignipediaglobal.pt/dicionario-global/deficiencia-direito-das-pessoas-com [accessed 9/5/2025].

[51] *The Framework Convention on Artificial Intelligence*, *op. cit.*, § 53.

[52] *Ibidem*, § 55.

technologies enhance rather than infringe upon their autonomy. The Drafters considered that referring to this concept in this Framework Convention is particularly appropriate in view of the capacity of artificial intelligence systems for imitation and manipulation."[53]

(b) Equality and non-discrimination: Algorithmic discrimination due to disability

Equality and non-discrimination are the classic principles of international human rights law. Discrimination based on AI activities is also included in the principled concept of equality and non-discrimination.[54] The principle of equality and non-discrimination under Article 10 implies that States guarantee "respect for equality," including gender equality, in the context of AI activities, as provided for in applicable international and national law. Not only does it respect equality, but it also requires States to address inequalities resulting from AI activities. The adoption of measures deemed appropriate to address these inequalities must "achieve impartial, fair, and equitable results, within the framework of their national and international human rights obligations."

Equally relevant, Article 17 of the Convention prohibits discrimination in the implementation of the Convention. The meaning of discrimination is identical to that established in applicable international law, for example, Articles 2 and 5 of the CRPD. The concept and scope of reasonable accommodations under Article 2 of the CRPD are fundamental in the context of the application of AI systems to ensure appropriate responses to the specific needs of people with disabilities. For example, AI systems must ensure compliance with rules regarding the design of accessible biometric systems or the provision of alternative means when technology failure poses a barrier that must be addressed to ensure that people with disabilities can benefit from the technology.

The definition of discrimination is broad and can include, for example, digital discrimination based on technology related to disadvantaged people with disabilities. As a result, the CRPD allows for a multiplicity of diverse concepts that define discrimination in Article 2 of the CRPD to keep pace with new social

---

[53] *Ibidem*.

[54] For example, Council of Europe, Christophe Lacroix, «Report on preventing discrimination caused by the use of artificial intelligence», PACE, Committee on Equality and Non-discrimination, Belgium, 29 Septembre, 2020; Tetyana (Tanya) Krupiy & Martin Scheinin, «Disability Discrimination in the Digital Realm: How the ICRPD Applies to Artificial Intelligence Decision-Making Processes and Helps in Determining the State of International Human Rights Law», *Human Rights Law Review*, Volume 23, Issue 3, September 2023, ngad019, https://doi.org/10.1093/hrlr/ngad019.

realities and, above all, technologies that ensure dynamic and evolving interpretation in changing circumstances. For example, digital discrimination is understood jointly under Articles 3, 4, 5, and 9 of the CRPD in conjunction with the FC-AI. For example, AI can replicate, exacerbating prejudice and discrimination against people with disabilities, due to errors or biases in datasets and algorithms that impact the rights of people with disabilities in various aspects of daily life.[55] The causes of algorithmic discrimination[56] can take both overt (i.e., prejudice, denial of opportunities) and subtle forms (i.e., negligence, ignorance, false assumptions, segregation, and social isolation through social barriers). States also have obligations to eliminate discrimination and provide avenues for accountability, as well as remedies for harm caused by discrimination. They must also effectively monitor the implementation of such measures to ensure that the measures are appropriate and effective in achieving the elimination of discrimination.

3.3.6. The rights of people with disabilities in light of the FC-AI and the CRPD

3.3.6.1. Meaning and scope of rights

Article 18 of the FC-AI refers to "the rights of persons with disabilities," applicable to the frameworks of rights enshrined in the FC-AI and, simultaneously, the CRPD as a whole. Just like the CRPD, which has two interconnected terms — "human rights and fundamental freedoms" — throughout the conventional text. Thus, the term "human rights" in the meaning of FC-AI "is intended to include "fundamental freedoms" in jurisdictions which draw a distinction between the two terms."[57]

We can say that the FC-AI addresses, so to speak, substantive and procedural rights in the area of AI. These are the "digital rights" related to the use of AI. In turn, the CRPD defines the set of human rights and fundamental freedoms in general. These are "human rights" with different ramifications. In other words, it addresses various substantive rights (e.g., the right to privacy) and procedural rights (e.g., the right to be consulted), as well as various obligations of public authorities under the same Convention. It is important to emphasize that States

---

[55] For example, European Union Agency for Fundamental Rights, *Bias in Algorithms – Artificial Intelligence and Discriminatio*n, Vienna, 2022.

[56] Thomas B. Nachbar, *Algorithmic Fairness, Algorithmic Discrimination*, 48 *Fla. St. U. L. Rev.*, 509, 2021, https://ir.law.fsu.edu/lr/vol48/iss2/5; Maarten Buyl; Christina Cociancig; Cristina Frattone & Nele Roekens, «Tackling Algorithmic Disability Discrimination in the Hiring Process: An Ethical, Legal and Technical Analysis», in *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*, June 21–24, 2022, Seoul, Republic of Korea, ACM, New York, NY, USA, 20 pages, https://doi.org/10.1145/3531146.3533169.

[57] *The Framework Convention on Artificial Intelligence*, *op. cit.*, § 5.

also have an obligation to ensure that "no person, organization, or private sector discriminates on the basis of disability" (Article 4(1)(e) CRPD). Furthermore, States have an obligation to promote the design and development of information technologies "at an early stage" (Article 9(2)(h) CRPD). This also covers issues related to AI systems, which cannot discriminate on the basis of disability, and States must adequately regulate the private sector regarding AI. We will now explain some substantive rights that are relevant in the context of AI. For example, the rights to privacy and data protection are equally fundamental to respect for human dignity.[58] AI systems must safeguard personal data about people with disabilities in any context, specifically regarding the content they create, the information collected about them, and what is inferred through algorithms.

Another example: The right to accessibility is enshrined in Article 9 of the CRPD. The right to accessibility is a fundamental element for considering AI. AI technologies and systems take into account the different specificities and particular needs of people with disabilities through technologies that facilitate daily life. Consequently, AI technologies must be designed in a way that is adaptable to these specific needs. States must take appropriate measures to guarantee people with disabilities access on an equal basis with others in any context.

Particularly in the case of deaf people and sign languages[59], in light of Article 18 of the FC-AI, the rights of deaf people take into account the specificities inherent to their respective sign languages, as well as the risks and vulnerabilities that may affect the exercise of language rights. Recently, both within the deaf community and in academia, there has been intense discussion about the use of AI for sign languages. In the context of technological advances, it is essential to ensure the rights of deaf people are fully respected and promoted, preventing new technologies from widening communication barriers. For example, deaf people should have the power to choose whether, when, and how to use AI technologies for communication, including the ability to opt for solutions involving human interpreters or assistive technologies, always respecting each person's individual freedom and personal autonomy in everyday life.

---

[58] *Cf.* Resolución aprobada por el Consejo de Derechos Humanos el 7 de octubre de 2021 48/4, *El derecho a la privacidad en la era digital*, Doc. A/HRC/RES/48/4.

[59] For example, Filipe Venade de Sousa, report published by EUD, "Sign Language in the Era of Artificial Intelligence", 2025; Louise Hickman; Alexa Hagerty & Rose Powell, «BSL Is Not For Sale: A Deaf-Led Approach to AI Procurement», MCTD Cambridge, 2025, DOI 10.17863/CAM.120927; «AI Interpreting Solutions Evaluation Toolkit: PART A: Organization, Implementation and Management», Developed by Stakeholders Advocating for Fair and Ethical AI in Interpreting Task Force (SAFE AI) & Coalition for Sign Language Equity in Technology (CoSET), August 2025.

3.3.6.2. Specific needs and vulnerabilities

Article 18 of the FC-AI introduces a particularly relevant obligation: to consider the *specific needs* and *vulnerabilities* of people with disabilities. However, the explanatory report does not explain what these terms mean. It is important to take into account the purpose and purpose of the FC-AI and the CRPD; specific needs and vulnerabilities relate to the conditions for exercising rights, not the intrinsic conditions of persons with disabilities themselves. In this sense, we must read the expression "in relation" as a prepositional phrase, not "of" as a preposition. In general, the notion of "specific needs" is inherent in the worthy and legitimate considerations in the everyday lives of people with disabilities, as well as appropriate conditions that must be ensured to facilitate the exercise of rights implied by AI.

The concept of vulnerability within the FC-AI is not legally defined.[60] We can take into account the example of the concept of "vulnerable" within EU law.[61] EU law often understands peoples with disabilities to fall under the concept of vulnerability in the sense that they may be susceptible to harm caused by AI systems that have a significant and unacceptable impact on them. Thus, specific vulnerabilities, in turn, refer to the increased risks that people with disabilities face in the context of AI.

According to the CRPD, disability is an "evolving concept" (preamble (e)) and "results from the interaction between people with disabilities and attitudinal and environmental barriers that impede their full and effective participation in society on an equal basis with others." (preamble (e) and Article 1) This understanding of vulnerability, which emphasizes the interaction between people with disabilities and their environment, essentially aligns with the definition of disability identified by the CRPD.

The vulnerability paradigm must be read in light of the social model of disability, which does not address the individual's particularities, but rather the social contexts that people with disabilities face in their daily lives. Vulnerability, as a legally relevant concept, encompasses situations in which people with disabilities need specific protection of their rights, identified as vulnerable, in order to gain access to certain forms of protection.

When interpreting the CRPD and the FC-AI, the concept of vulnerability should not be confused with the stigmatization and stereotyping that such

---

[60] European Commission, «ETHICS GUIDELINES FOR TRUSTWORTHY AI», High-Level Expert Group on Artificial Intelligence, Document made public on 8 April 2019.

[61] European Commission, «Approval of the content of the draft Communication from the Commission - Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)»; «Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)».

concepts generally imply, resulting from discriminatory or stereotypical patterns incompatible with the purposes of the CRPD. It should be addressed in negative terms regarding the enjoyment and exercise of rights and freedoms that may hinder adequate access in terms of AI. It should not be addressed in negative terms regarding the identification of disability *per se*, within the logic of the welfare approach, and should ensure approaches compatible with the purposes of the CRPD and the FC-AI. In this logic, AI systems and technologies should address vulnerabilities as a result of the social environment that faces barriers rather than resulting from a people with disability. Thus, in the context of AI use, vulnerability is extrinsic to the condition of people with disabilities who enjoy and exercise their rights within the framework of the social barriers they face. Distinguishing between different sources of vulnerability, vulnerability is essentially caused by the interaction between people with disabilities and their social environment, resulting from situational circumstances arising or exacerbated by various environmental, behavioral, and social factors within AI systems.

3.3.6.3. Risk assessment in the context of the rights of people with disabilities

Article 16 establishes the meaning and scope of the classification and management of risks and impacts resulting from the use of AI, taking into account the "real and potential" implications for human rights. Such measures will be graduated and differentiated, as appropriate, and will take into account the context and intended use of artificial intelligence systems, particularly with regard to risks to human rights.[62]

The risks and vulnerabilities inherent to the rights of people with disabilities must correspond to the contexts of the social model of disability, or rather, the human rights model enshrined in the CRPD. People with disabilities, like everyone else, have vulnerabilities inherent in the social contexts they face in their daily lives, with different ramifications that lead to discrimination. AI systems can exacerbate discrimination against people with disabilities, and these systems necessarily operate based on data provided by various sources that can reproduce the same discriminatory patterns. Therefore, people with disabilities face a greater risk of violations of their fundamental rights and freedoms.

The CRPD and FC-AI require government authorities to take measures to address and reduce risks and vulnerabilities in accordance with conventional standards, and, above all, to adopt appropriate approaches that reflect human dignity, equality, and non-discrimination. It is important to emphasize that the emphasis is on the phenomenon of discrimination and not on individual characteristics.

[62] *The Framework Convention on Artificial Intelligence*, *op. cit.*, § 108.

Taking into account the principles and standards of the CRPD and the FC-AI, States must adopt necessary measures related to risk prevention, mitigation, and monitoring, taking into account the needs and vulnerabilities of peoples with disabilities, to establish and identify parameters for assessing the impact and benefits useful for the enjoyment and exercise of fundamental rights and freedoms of peoples with disabilities. Such impact assessments must identify impacts on the rights of people with disabilities in all situations that may undermine the enjoyment and exercise of these rights.

The FC-AI requires States to adopt appropriate measures to identify, assess, prevent, and mitigate, *ex ante* and, as appropriate, iteratively throughout the lifecycle of the AI system, the relevant risks and potential impacts to human rights, democracy, and the rule of law, following and enabling the development of a methodology with concrete and objective criteria for such assessments.[63]

Risk assessment requires "graduated and differentiated" measures, whenever appropriate, which must meet several assumptions for the adoption of these measures. These measures take into account[64]:

i. the context and intended use of the systems developed by AI;

ii. the severity and likelihood of potential impacts;

iii. the views of relevant stakeholders, where appropriate;

iv. interactive application throughout the activities carried out within the scope of AI;

v. monitoring of risks and negative impacts on human rights, democracy, and the rule of law;

vi. documentation of risks, actual and potential impacts, and the risk management approach;

vii. requirement for prior testing of AI systems before they are made available for first use and when they undergo significant modifications.

In addition to these assumptions, the adoption or maintenance of measures that ensure "adverse"[65] impacts are "adequately address"[66] must also be documented and taken into account in the development of relevant risk management measures. And, likewise, States assess the need for a "moratorium, prohibition or other appropriate measures" in relation to "certain uses" of AI systems

[63] *Ibidem*, § 105.

[64] *Ibidem*, § 107.

[65] *Ibidem*, §§ 109-110.

[66] *Ibidem*, § 110.

when they consider that such uses are "incompatible" with respect for human rights, the functioning of democracy or the rule of law.[67]

4. Conclusions

The legal frameworks regarding the use of AI in the context of the rights of people with disabilities, in light of the FC-AI and the CRPD, are relevant and relevant today. These Conventions respond, from the perspective of better rights protection, to the new frontiers of the protection of the rights of people with disabilities in the field of AI.

As a result, AI Law, as a specialized area of Technology Law, defines, at a dogmatic level, the rights related to the use of AI, and any technology created and developed today, which must ensure respect for the fundamental rights and freedoms of people with disabilities. The use of AI or technological solutions meets the legitimate expectations of people with disabilities, and not the other way around. Thus, Inclusive Digital Law consists of guaranteeing the right to self-determination – in a multidimensional sense – in the digital and technological space, which relates to the right to access the digital world, preserving the right to exist in a digital world, e.g., AI technologies and systems that offer various services to the public.

The right to self-determination is rooted in personal autonomy and the individual freedom to make choices, *as per* Article 3 of the CRPD. In other words, people with disabilities have the right to decide when and within what limits technological solutions or AI systems can be used to meet their specific needs.

The perspective that the right to self-determination in the inclusive digital realm requires that any and all technological approaches be accessible and available to anyone who can benefit from and exercise personal self-determination in the digital space. This constitutes a dimension of the right to accessibility, benefiting, as such, from the protection granted to it in Article 9 of the CRPD. Thus, the right to self-determination in the inclusive digital realm is characterized as the right of people with disabilities to define and select what is best for their legitimate interests within the framework of the digital and technological solutions available.

[67] *Ibidem*, § 111.

ARTIFICIAL INTELLIGENCE AT THE SERVICE OF SUSTAINABLE DEVELOPMENT AND CORPORATE GOVERNANCE: BRIEF NOTES ON LEGAL ASPECTS OF ESG GOALS AND SDGS

*Hana Horak**

Abstract: Artificial Intelligence (AI) is increasingly influencing both sustainability strategies and corporate governance frameworks. This paper examines how AI can serve at the service of the realisation of Environmental, Social, and Governance (ESG) goals and the United Nations Sustainable Development Goals (SDGs), while reinforcing accountability, transparency, and oversight functions within corporate governance. The legal and ethical implications of AI usage—particularly issues of algorithmic bias, data protection, and decision-making liability—are analysed in light of the European Union's Artificial Intelligence Act (2024) and national legal frameworks such as the Croatian Companies Act. The paper argues that corporate governance must evolve to accommodate the

* PhD in Law; Jean Monnet Chair and Full Professor (Tenured), Department of Business Law, Faculty of Economics and Business, University of Zagreb, Croatia; President of the ICC Croatia Business Integrity Commission; Editor-in-Chief of the InterEULawEast Journal. Orcid: 0000-0002-8985-2589. E-mail: hhorak@net.efzg.hr.

transformative potential of AI by integrating transparent and explainable (rather than merely "transparent") AI systems, ensuring traceability and accountability in algorithmic decision-making processes. Furthermore, the reinterpretation of the duty of care of management and supervisory boards is discussed as a legal response to the challenges of algorithmic governance and automation risk.

Finally, the concept of corporate purpose is examined as an ethical and normative corrective in the digital era, positioning AI as a potential enabler of sustainable corporate decision-making consistent with long-term stakeholder interests. The analysis highlights the need for greater alignment between AI governance, ESG compliance, and existing legal duties of directors, emphasising that AI can meaningfully contribute to corporate sustainability only within robust regulatory and ethical frameworks.

Keywords: Artificial Intelligence; ESG; Corporate Governance.

Resumo: A Inteligência Artificial (IA) está a influenciar cada vez mais as estratégias de sustentabilidade bem como as estruturas de governação corporativa. Este artigo analisa como a IA pode estar ao serviço da concretização dos objetivos ambientais, sociais e de governação (ESG) e dos Objetivos de Desenvolvimento Sustentável (ODS) das Nações Unidas, reforçando simultaneamente as funções de responsabilização, transparência e supervisão no âmbito da governação corporativa. As implicações legais e éticas do uso da IA – nomeadamente questões de viés algorítmico, proteção de dados e responsabilidade pela tomada de decisões – são analisadas à luz da Lei de Inteligência Artificial da União Europeia (2024) e de estruturas legais nacionais, como a Lei das Sociedades da Croácia.

O artigo defende que a governação corporativa deve evoluir para acomodar o potencial transformador da IA, integrando sistemas de IA transparentes e explicáveis (em vez de meramente "transparentes") que garantam a rastreabilidade e a responsabilidade nos processos de tomada de decisão algorítmica. Além disso, a reinterpretação do dever de diligência dos conselhos de administração e de supervisão é discutida como uma resposta jurídica aos desafios da governação algorítmica e do risco de automação.

Por fim, o conceito de objetivo corporativo é examinado como uma correção ética e normativa na era digital, posicionando a IA como um potencial facilitador da tomada de decisões corporativas sustentáveis, consistentes com os interesses de longo prazo das partes interessadas. A análise destaca a necessidade de um maior alinhamento entre a governação da IA, a conformidade com os critérios ESG e os deveres legais existentes dos diretores, enfatizando que a IA só pode contribuir significativamente para a sustentabilidade corporativa dentro de estruturas regulatórias e éticas robustas.

Palavras-chave: Inteligência Artificial; ESG: Governação Corporativa.

1. Introduction

Artificial Intelligence (AI) is increasingly recognised as a transformative technology for advancing global sustainability targets, particularly the United Nations Sustainable Development Goals (SDGs) and Environmental, Social, and Governance (ESG)[1] criteria. By optimising resource allocation, enabling data-driven decisions, and fostering innovative solutions to challenges such as climate change and social inequality, AI offers significant potential for sustainable development. However, AI integration also poses complex legal and ethical questions, including issues of data privacy, algorithmic bias, and accountability.[2] Recent regulations—most notably the European Union's AI Act[3] and international initiatives such as the United Nations' resolution on *safe, secure, and trustworthy AI*[4]—underscore the need for a balanced approach that upholds fundamental rights while fostering technological innovation.

This paper examines the interplay between AI, ESG, and the SDGs from a legal perspective, highlighting the roles of policymakers, businesses, and legal practitioners in shaping responsible AI adoption. Although technological advancement brings substantial benefits in terms of efficiency and accessibility of information, it also gives rise to intricate legal and ethical challenges concerning the responsibilities of management and supervisory boards, the transparency and explainability of algorithmic decision-making, and the establishment of new supervisory mechanisms.

Special emphasis is placed on: (1) the responsibility of members of management boards, supervisory boards, and boards of directors; (2) the establishment of explainable and supervised algorithmic systems (explainable AI); and (3) the inclusion of the purpose of the company (corporate purpose) as an ethical corrective in the digital age.

---

[1] François Giguère, Redefining Industrial Practices: AI's Ethical Integration into Tomorrow's Markets, 2024, DOI:10.13140/RG.2.2.18614.51523; Massimo Regona, Tan Yigitcanlar, Carol Hon, & Melissa Teo, «Artificial Intelligence and Sustainable Development Goals: Systematic Literature Review of the Construction Industry», Sustainable Cities and Society 108, 2024, p. 105499, https://doi.org/10.1016/j.scs.2024.105499, [consult. at 17/07/2025].

[2] UNESCO, Recommendations on the Ethics of Artificial Intelligence, Paris, UNESCO Publishing, 2023.

[3] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139, and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797, and (EU) 2020/1828 (Artificial Intelligence Act), Official Journal of the European Union, L, 2024, pp. 1–320.

[4] UN General Assembly, Resolution A/78/L.49 on Safe, Secure, and Trustworthy AI, New York, United Nations, 2024.

Drawing on regulatory frameworks, best practices, and relevant scientific literature on ESG and SDGs, this paper argues for robust governance structures that integrate AI into sustainability efforts without compromising ethical principles or human rights.

The rapid deployment of AI across diverse industries, from healthcare to finance and transportation, has generated unprecedented opportunities to address sustainability challenges.[5] Simultaneously, the international community has intensified its commitment to responsible business practices through frameworks such as the SDGs and ESG criteria.[6]/[7]

As companies leverage AI to optimise processes, enhance decision-making, and design innovative solutions for climate action and social development, they also face intricate legal, ethical, and operational considerations.[8]

AI has thus become a core element of contemporary corporate responsibility agendas. Modern boards and executives increasingly recognise that responsible AI deployment extends beyond regulatory compliance; it influences public trust and contributes to long-term corporate value. This recognition raises critical governance issues, including algorithmic bias, data protection, and the environmental implications of data-intensive processes.[9]

Reflecting these concerns, legislation such as the EU's AI Act, initiated in 2024 and entering into force over the following two years, seeks to mitigate the risks associated with AI while promoting innovation. On the global level, the United Nations resolution A/78/L.49, adopted in March 2024, reinforces the principle of "safe, secure, and trustworthy AI", confirming international consensus on integrating AI into sustainable development.

This paper synthesises these developments, focusing on how AI can be employed to achieve SDG-related targets and ESG standards within evolving legal frameworks. It also addresses operational challenges such as workforce training and infrastructure modernisation.

---

[5] François Giguère, Redefining Industrial Practices…, op. cit.; Massimo Regona et al., «Artificial Intelligence and Sustainable Development Goals...», op. cit.

[6] United Nations, Transforming Our World: The 2030 Agenda for Sustainable Development, New York, United Nations, 2015, https://docs.un.org/en/A/RES/70/1.

[7] World Economic Forum, Global Technologies in Sustainability: AI-Driven Approaches, Geneva, WEF Publications, 2022, https://reports.weforum.org/docs/WEF_Artificial_Intelligence_for_Efficiency_Sustainability_and_Inclusivity_in_TradeTech_2025.pdf.

[8] OECD, Recommendation of the Council on Artificial Intelligence, Paris, OECD Publishing, 2023, https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449.

[9] OECD, «Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint», OECD Digital Economy Papers, No. 341, OECD Publishing, Paris, 2022, https://doi.org/10.1787/7babf571-en.

That businesses must navigate to integrate AI effectively. The legislative and regulatory framework of the European Union, particularly through the AI Act (2024), sets fundamental rules for the development and use of AI systems but does not directly regulate corporate governance. National legislation, such as the Companies Act of the Republic of Croatia, similarly fails to provide explicit guidance on delegating decision-making to AI systems, relying instead on the traditional interpretation of the principle of due diligence of board members.

In this context, the integration of AI into management processes necessitates a reinterpretation of existing legal concepts and the creation of specialised regulatory and self-regulatory mechanisms. The fundamental principles of corporate governance—responsibility, transparency, and supervision—can be preserved and even strengthened within an increasingly complex algorithmic environment.

Special emphasis is therefore placed on maintaining the accountability of members of management boards, supervisory boards, and boards of directors; ensuring the establishment of explainable and auditable algorithmic systems (explainable AI); and reaffirming corporate purpose as an ethical and normative corrective in the governance of digital transformation.[10]

2. The AI–sustainability nexus

2.1. *AI as a catalyst for sustainability*

AI technologies hold considerable potential to address the core concerns of the SDGs—namely climate change, social inequality, and economic disparities. AI-driven predictive analytics, for instance, can optimise energy consumption and thereby reduce carbon emissions in industrial processes. Likewise, machine learning algorithms can identify inefficiencies in global supply chains, enabling businesses to cut resource usage and minimise waste.

In the financial sector, AI-powered credit-scoring models can extend access to financial services for underserved populations, thereby contributing to social inclusion and the reduction of inequality.[11]

---

[10] Global AI standards often group transparency and explainability together because they are interconnected terms. Transparency answers the question "what happened" in the AI system, while explainability addresses "how" a decision was made using AI. See also European Law Institute, Guiding Principles for Automated Decision-Making in the EU, Vienna, ELI, 2022, https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Guiding_Principles_for_Automated_Decision-Making_in_the_EU.pdf.

[11] OECD, «Regulatory approaches to Artificial Intelligence in finance», OECD Artificial Intelligence Papers, No. 24, OECD Publishing, Paris, 2024, https://doi.org/10.1787/f1498c02-en. https://www.undp.org/sites/g/files/zskgke326/files/2025-05/hdr_2025_overview.pdf.

In line with ESG criteria, AI tools facilitate real-time data collection, monitoring, and reporting on environmental, social, and governance metrics. Automated systems can detect early indicators of fraud, corruption, or unethical labour practices, thus strengthening corporate governance and enhancing accountability. When integrated effectively, these capabilities transform sustainability from a peripheral compliance issue into a central strategic function, enabling companies to align operations with measurable sustainability goals.

Nevertheless, the deployment of AI in sustainability initiatives introduces new challenges that must be addressed through sound legal, ethical, and technical frameworks.

AI-driven data processing requires substantial computational power, which may significantly increase an organisation's carbon footprint, particularly in data-intensive sectors such as cloud computing and logistics.[12]

Furthermore, the reliance on extensive datasets raises questions about privacy, proportionality, and data protection, especially in light of the General Data Protection Regulation (GDPR).

Algorithmic bias represents another critical concern. When training data lack diversity or contain embedded societal biases, AI systems can produce discriminatory outcomes that undermine social equity objectives and ESG commitments. Such bias can also expose companies to legal liability and reputational damage, particularly where AI-driven decisions, such as automated hiring or credit evaluation, have direct impacts on individuals. For example, an AI-powered recruitment platform that unintentionally discriminates against certain demographic groups contradicts a company's social responsibility objectives and may breach both equality and data protection laws.

Regulators are increasingly attentive to these risks, with emerging governance frameworks emphasising consumer protection, privacy, and fairness as fundamental principles of responsible AI use. Effective mitigation strategies therefore require a combination of legal due diligence and operational safeguards, including regular algorithmic audits, explainability assessments, and human oversight of automated decisions. Organisational measures such as

[12] Peter Gailhofer; Anke Herold; Jan Peter Schemmel; Cara-Sophie Scherf; Cristina Urrutia; Andreas R. Köhler & Sibylle Braungardt, «The role of Artificial Intelligence in the European Green Deal», Directorate-General for Internal Policies, 2021, https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662906/IPOL_STU%282021%29662906_EN.pdf.

workforce training, interdisciplinary compliance committees, and transparent disclosure mechanisms further enhance accountability.[13]

The integration of AI into sustainability strategies thus demands a multidimensional governance approach—combining technological innovation with strong legal and ethical controls. Properly designed and supervised, AI can act as a genuine catalyst for sustainability, provided that transparency, explainability, and accountability remain at the core of its deployment.

3. Brief notes on corporate governance, ESG, and AI

3.1. *ESG and SDGs: A legal and ethical context*

ESG criteria have become standard benchmarks for assessing a company's environmental stewardship, social responsibility, and governance practices.[14] Meanwhile, the SDGs, adopted by the United Nations in 2015, extend these objectives to include the eradication of poverty, the promotion of quality education, and the mitigation of climate change on a global scale. Together, ESG and the SDGs provide a comprehensive framework for evaluating the sustainability performance and social legitimacy of corporate activities.

Scientific studies emphasise the interdependence of sustainability, financial performance, and stakeholder trust. For instance, Eccles, Ioannou, and Serafeim[15] demonstrate that companies with robust sustainability policies tend to exhibit superior organisational processes and long-term performance, while Bebbington and Unerman underscore the importance of transparent corporate reporting in advancing the SDGs. These findings reinforce the notion that sustainability is not merely a normative ideal but a material determinant of corporate success and accountability.[16]

---

[13] OECD, Recommendation of the Council on Artificial Intelligence, 2025, OECD/LEGAL/0449, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#mainText.

[14] Jan Bebbington; Jeffrey Unerman, «Achieving the United Nations Sustainable Development Goals: An Enabling Role for Accounting Research», Accounting, Auditing & Accountability Journal 31, no. 1, 2018, pp. 2–24.

[15] Robert G. Eccles, Ioannis Ioannou, & George Serafeim, «The Impact of Corporate Sustainability on Organizational Processes and Performance», Management Science 60, no. 11, 2014, pp. 2835–2857, https://doi.org/10.1287/mnsc.2014.1984.

[16] OECD, Global Corporate Sustainability Report 2024, OECD Publishing, Paris, 2024, https://doi.org/10.1787/8416b635-en; OECD, «ESG ratings and climate transition: An assessment of the alignment of E pillar scores and metrics», OECD Business and Finance Policy Papers, No. 6, OECD Publishing, Paris, 2022, https://doi.org/10.1787/2fa21143-en; OECD, Behind ESG Ratings, OECD Publishing, Paris, 2025, https://doi.org/10.1787/3f055f0c-en.

As companies increasingly align with these frameworks, Artificial Intelligence (AI) has emerged as a strategic tool for achieving and reporting on ESG objectives. In the environmental domain, AI-driven models enable more accurate forecasting of resource consumption and contribute to the reduction of greenhouse gas emissions through enhanced energy efficiency and waste management. In the social dimension, companies are leveraging AI to strengthen diversity, equity, and inclusion initiatives, support community welfare programmes, and improve workplace safety through predictive analytics. Regarding governance, automated oversight systems powered by AI enhance compliance and ethical standards by facilitating the early detection of anomalies, fraud, and other forms of misconduct.

Implementing AI responsibly across these areas requires cross-functional collaboration. Sustainability teams must define environmental and social objectives consistent with ESG and SDG indicators, while technical teams ensure the reliability, transparency, and ethical design of AI solutions. Legal departments play a crucial role in embedding due diligence, data protection, and accountability mechanisms into AI governance structures. Such interdisciplinary cooperation ensures that AI contributes not only to operational efficiency but also to the integrity, traceability, and credibility of sustainability reporting.

3.2. *AI governance and board responsibility*

The responsibility of members of the management board and the supervisory board in the use of AI includes the application of the principle of duty of care, which encompasses the obligation of informed, diligent, and loyal behaviour in the interest of the company.[17]

In the European context, this is reflected in Directive (EU) 2017/1132 on certain aspects of company law, as well as in national legislation—such as Article 252 of the Companies Act of the Republic of Croatia, which regulates the liability of board members.[18]

A key question arises: Is the use of algorithmic recommendations, when accompanied by a proper understanding of their limitations, consistent with the expected standard of care for board members? According to an increasingly accepted view in both theory and practice, the adoption of advanced technologies can demonstrate prudent management. Brand argues that longstanding direc-

[17] Hana Horak; Lucia Ana Tomić, «Directors, Information, Remuneration and Risk Management», European Company Case Law (ECCL), 2, 2023, pp. 149–162, https://doi.org/10.5771/2752-177X-2023-2-149.

[18] Official Gazette 111/93, 34/99, 121/99, 52/00, 118/03, 107/07, 146/08, 137/09, 125/11, 152/11, 111/12, 68/13, 110/15, 40/19, 34/22, 114/22, 18/23, 130/23, 136/24.

tors' duties (care, diligence, loyalty) can and should be read in a way that supports AI governance objectives; maps oversight obligations to AI risk.[19]

However, it does not relieve directors of their overarching duty to supervise and evaluate these systems critically. The "traditional" duties can and must be extended to use of AI bearing in mind new director's competencies.[20]

Emerging aspects of responsibility now include new competencies. In particular, "literacy in the field of artificial intelligence" refers to the skills, knowledge, and understanding that enable suppliers, deployers, and affected persons, within the framework of the AI Act to introduce and manage AI systems in an informed manner, with awareness of their potential risks, benefits, and harms. For corporate boards, this implies that both management and supervisory board members must acquire a baseline understanding of AI's functioning and limitations in order to exercise meaningful oversight.[21]

Supervisory board members increasingly bear responsibility for overseeing algorithmic systems used in strategic or operational functions, particularly where such systems influence corporate decision-making or stakeholder outcomes. In this context, the AI Act has a far-reaching impact on corporate governance. Its key provisions establish a comprehensive framework for the regulation of artificial intelligence systems. First, the Act introduces the categorisation of AI systems according to defined risk levels (Articles 6–9), distinguishing between unacceptable, high, limited, and minimal-risk applications. Second, it imposes compliance assessment and documentation obligations for high-risk systems (Articles 16–29), requiring providers and deployers to implement conformity procedures and maintain detailed technical documentation. Third, it places explicit responsibility on the companies for the implementation, management, and continuous monitoring of AI systems, thereby ensuring that accountability and oversight remain central to corporate governance practices.

Corporate boards and senior executives thus increasingly view AI governance as an integral component of risk management and compliance frame-

---

[19] Vivianne Brand, «Directors' Duties and AI Regulation», Griffith Law Review, 33(4), 2024, pp. 324–354, https://doi.org/10.1080/10383441.2024.2422221.

[20] See more on duty of care in: Klaus J. Hopt, «Directors' Duties and Shareholders' Rights in the European Union», ECGI Law Working Paper, 2016, https://www.ecgi.global/sites/default/files/working_papers/documents/SSRN-id2749237.pdf.

[21] The concept of AI literacy mentioned in Article 4 of the AI Act relies on the definition of the term given in Article 3(56) of the AI Act, according to which: "AI literacy' means skills, knowledge and understanding that allow providers, deployers and affected persons, taking into account their respective rights and obligations in the context of this Regulation, to make an informed deployment of AI systems, as well as to gain awareness about the opportunities and risks of AI and possible harm it can cause.", https://digital-strategy.ec.europa.eu/en/policies/ai-talent-skills-and-literacy.

works. According to international principles such as the OECD's AI Principles, companies deploying AI should maintain transparency, accountability, and fairness in their oversight processes. This entails evaluating AI systems for potential harms related to privacy, non-discrimination, and consumer protection in order to ensure that algorithmic outcomes comply with fundamental rights and ethical standards. It also requires adherence to data collection and processing rules consistent with the GDPR, thereby safeguarding the integrity and lawfulness of data use within corporate operations. Furthermore, companies should employ staff with sufficient AI literacy to interpret algorithmic outputs, identify potential errors or biases, and intervene when necessary to prevent adverse consequences. Finally, boards are encouraged to establish cross-disciplinary committees composed of legal, technical, and ethical experts to oversee responsible AI adoption and to ensure that decision-making processes remain transparent, accountable, and aligned with corporate sustainability objectives.

These governance structures ensure that AI activities support a company's sustainability and ESG objectives without compromising fundamental rights or the principles of lawful, fair, and transparent corporate conduct. This approach aligns with the obligations set out in the Corporate Sustainability Due Diligence Directive (CSDDD), which emphasises the integration of due diligence processes into corporate strategies and risk management systems. By embedding AI oversight within the broader due diligence framework, companies can strengthen accountability and demonstrate compliance with evolving European standards on sustainable and responsible business conduct.[22]

4. The legal and ethical dimensions of AI in ESG

4.1. *Regulatory frameworks: The EU's AI Act*

A cornerstone of the emerging regulatory landscape, EU AI Act establishes a comprehensive, risk-based framework for the development, placement, and use of AI systems in the internal market. It classifies AI applications into four categories: unacceptable, high-risk, limited-risk, and low or minimal-risk, imposing correspondingly stricter requirements for high-risk systems.

Key obligations under the Act include the conduct of risk assessments, conformity and compliance checks, and the adoption of transparency, documen-

[22] Directive (EU) 2024/1760 of the European Parliament and of the Council of 24 July 2024 on Corporate Sustainability Due Diligence and amending Directive (EU) 2019/1937 (Corporate Sustainability Due Diligence Directive), Official Journal of the European Union L, 2024, pp. 1–75.

tation, and oversight measures. Given its extraterritorial reach, any provider or deployer placing an AI system on the EU market is subject to these obligations, irrespective of geographic location.[23]

For companies utilising AI to advance ESG and SDG targets, compliance with the AI Act represents both a legal and reputational imperative. High-risk AI systems, such as those applied in lending decisions, human-resources management, or critical infrastructure, require systematic documentation, robust internal controls, and continuous staff training. Failure to meet these obligations may expose organisations to administrative penalties, civil liability, and significant reputational harm, thereby undermining their broader sustainability and governance commitments.

By embedding the principles of accountability, transparency, and due diligence within corporate AI governance frameworks, companies can align with the EU's evolving regulatory standards while reinforcing stakeholder trust. This coherence between AI oversight and sustainable business conduct is further reflected in the OECD Recommendation of the Council on Artificial Intelligence (2023), which promotes the principles of transparency, accountability, and human-centred development of AI, and in the Corporate Sustainability Due Diligence Directive which requires companies to integrate due-diligence processes addressing actual and potential adverse impacts of their operations, including those arising from the use of advanced digital technologies. Together, these instruments reinforce the expectation that responsible AI governance forms part of a company's overarching sustainability and risk-management framework.

4.2. *The United Nations Resolution and international law*

At the international level, the United Nations General Assembly Resolution on "Safe, Secure and Trustworthy Artificial Intelligence" (A/78/L.49), adopted in March 2024, articulates a global framework of shared principles for the responsible design, development, and use of AI.[24] Although not legally binding, the Resolution represents a significant normative milestone in the consolidation of international expectations concerning AI governance. It calls upon States and non-State actors alike to ensure that AI systems are developed and deployed in accordance with international human rights law, democratic values, and the rule

[23] Thom-Okoroh Biodomoye, «The EU AI Act and Its Extraterritorial Reach: Implications for African AI Providers», June 14, 2025, https://doi.org/10.2139/ssrn.5295041.

[24] UN General Assembly, Resolution on "Safe, Secure and Trustworthy Artificial Intelligence" (A/78/L.49), adopted 21 March 2024, Official Records of the General Assembly, Seventy-eighth Session, Agenda Item 28 (b), New York, United Nations, 2024.

of law, thereby emphasising that technological innovation must not come at the expense of fundamental freedoms.

The Resolution recognises AI's potential to serve as an accelerator for the implementation of the United Nations SDGs, particularly in areas such as climate action, education, and healthcare. However, it simultaneously underscores the necessity of risk mitigation and accountability mechanisms to prevent the misuse of AI in ways that could entrench inequality or infringe individual rights. In this respect, it establishes a moral and policy foundation for integrating human rights due diligence into AI governance, anticipating developments later codified within EU legal framework such as the Corporate Sustainability Due Diligence Directive.

Furthermore, the Resolution urges private entities to align their practices with the UN Guiding Principles on Business and Human Rights (2011), reaffirming the tripartite structure of State duty to protect, corporate responsibility to respect, and access to remedy.[25] This expectation extends explicitly to the deployment of AI technologies, where algorithmic decisions may affect individuals' rights to privacy, equality, and non-discrimination

In parallel, complementary international instruments such as the OECD Guidelines for Multinational Enterprises on Responsible Business Conduct (2023)[26] and the UNESCO Recommendation on the Ethics of Artificial Intelligence (2023)[27] reinforce the same normative trajectory. They collectively promote transparency, explainability, inclusiveness, and environmental sustainability as guiding principles for AI development. Taken together, these soft-law instruments contribute to the formation of a transnational governance ecosystem, bridging voluntary corporate ethics with emerging legal obligations under regional frameworks like the EU's AI Act and the CSDDD.

Adherence to these principles not only strengthens a company's global legitimacy and mitigates cross-border legal risks but also integrates AI governance within a rights-based approach to sustainability. In doing so, it positions responsible AI deployment as an essential component of modern corporate citizenship and sustainable value creation.

---

[25] United Nations, Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework, HR/PUB/11/04, New York and Geneva, United Nations, Office of the High Commissioner for Human Rights, 2011.

[26] OECD, OECD Guidelines for Multinational Enterprises on Responsible Business Conduct, Paris, OECD Publishing, 2023.

[27] UNESCO, Recommendation on the Ethics of Artificial Intelligence, Paris, UNESCO Publishing, 2023.

4.3. *Ethical challenges: Bias, accountability, and transparency*

Ethical considerations are central to embedding artificial intelligence within ESG strategies and sustainable corporate governance. Algorithmic bias remains one of the most pressing ethical risks, as it can inadvertently reproduce or amplify existing patterns of discrimination, thereby conflicting with the social-equity and inclusion objectives that ESG frameworks seek to advance. Determining accountability for AI-driven outcomes, particularly when harm results from autonomous or semi-autonomous decision-making, raises complex legal questions at the intersection of corporate, data-protection, and tort law.

It is essential to distinguish between transparency and explainability, two concepts often used interchangeably but representing distinct dimensions of algorithmic governance. *Transparency* concerns the openness and accessibility of information about how an AI system is designed, trained, and deployed, thus enabling regulatory scrutiny and public oversight. *Explainability*, by contrast, refers to the capacity to interpret and understand the reasoning or logic underlying an algorithm's outputs, allowing decision-makers and affected individuals to assess whether outcomes are fair, lawful, and ethically acceptable.

This distinction has been explicitly recognised in the European Commission's Ethics Guidelines for Trustworthy AI (2019)[28], which identify transparency as a foundational requirement encompassing traceability, explainability, and communication. Likewise, Wachter, Mittelstadt and Russell emphasise that transparency concerns the disclosure of system information, whereas explainability entails meaningful interpretability by human agents.[29]

Also, they argue that many advanced AI models, particularly those relying on deep-learning architectures, operate as "black boxes", making it difficult for companies to provide intelligible explanations for algorithmic decisions. Such opacity can erode stakeholder trust, obstruct effective regulatory supervision, and diminish the perceived legitimacy of corporate ESG efforts.

Addressing these ethical and legal challenges requires organisation-wide measures. Companies should conduct regular algorithmic audits and impact assessments to detect and mitigate risks of bias, discrimination, or other harmful effects. They must also maintain comprehensive documentation of model development, data sources, and validation processes to demonstrate due diligence and enable accountability. Contractual frameworks should clearly allocate liability

---

[28] European Commission, Ethics Guidelines for Trustworthy AI, Brussels, High-Level Expert Group on Artificial Intelligence, 2019.

[29] Sandra Wachter, Brent Mittelstadt, & Chris Russell, «Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI», Computer Law & Security Review 41, 2021, p. 105567, https://doi.org/10.1016/j.clsr.2021.105567.

between AI providers, deployers, and users, ensuring that governance responsibilities remain traceable throughout the AI lifecycle.

Legal practitioners play a pivotal role in ensuring that human-rights due diligence encompasses all stages of AI deployment, from design and procurement to operation and post-market monitoring, in accordance with the Corporate Sustainability Due Diligence Directive and the UN Guiding Principles on Business and Human Rights (2011). Embedding these standards within corporate practice not only enhances legal compliance but also strengthens ethical legitimacy and stakeholder trust, ensuring that AI contributes positively to sustainable and socially responsible value creation.

5. Emerging regulatory approaches and best practices

5.1. *National and regional regulations*

Beyond the European Union, several jurisdictions have adopted or proposed distinct regulatory approaches to AI governance, creating a fragmented global landscape that complicates compliance for multinational companies[30] seeking to fulfil ESG obligations.

The United States, for instance, has taken a predominantly sector-based approach, focusing on data-protection and consumer-protection statutes[31], while China's regulatory model combines state-driven oversight with ethical principles emphasising social harmony and public security.[32]

The United Kingdom's Pro-Innovation AI Regulation Framework (2023)[33] promotes a principles-based, regulator-specific approach rather than a single statutory instrument. These divergent models underscore the difficulty of harmonising AI governance at the global level.

---

[30] OECD, State of Implementation of the OECD AI Principles: Insights from National AI Policies, Paris, OECD Publishing, 2021, https://www.oecd.org/publications/state-of-implementation-of-the-oecd-ai-principles-1cd40c44-en.htm [consult. at 17/07/2025].

[31] «The U.S. Approach to Artificial Intelligence – Federal Laws, Policies & Practices» (European Parliament study), 2024, https://www.europarl.europa.eu/RegData/etudes/ATAG/2024/757605/EPRS_ATA(2024)757605_EN.pdf.

[32] «Balancing innovation and regulation: comparing China's AI regulations with the EU AI Act» (AWA article), 2024, https://www.awa.com/awapoint/balancing-innovation-and-regulation-comparing-chinas-ai-regulations-with-the-eu-ai-act/.

[33] «AI regulation: A pro-innovation approach» (UK Government publication), 2023, https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach.

Nevertheless, multilateral forums, such as the G20, the OECD, and the United Nations Educational, Scientific and Cultural Organization (UNESCO), have advanced consensus around shared principles of responsible AI. The OECD Principles on Artificial Intelligence (2019), later endorsed by the G20, constitute the first intergovernmental standard promoting values-based AI that is inclusive, transparent, robust, and accountable. These principles emphasise: (1) human-centred values and fairness; (2) transparency and explainability; (3) robustness, security, and safety; (4) accountability; and (5) inclusive growth and sustainable development. Together with the UNESCO Recommendation on the Ethics of Artificial Intelligence (2023), they form the foundation of a converging international soft-law framework that increasingly informs national legislation and corporate practice.

In this evolving regulatory context, many companies have begun to institutionalise AI governance by establishing interdisciplinary AI task forces comprising data scientists, legal practitioners, compliance officers, and ethicists. Such cross-functional collaboration enables holistic risk assessments and early identification of potential legal and ethical pitfalls. Drawing on normative frameworks such as the OECD AI Principles and the Corporate Sustainability Due Diligence Directive, companies are developing internal AI ethics policies that address fairness, privacy, human oversight, and security. Integrating these commitments into corporate codes of conduct and sustainability governance frameworks promotes consistency across operational practices and strengthens accountability mechanisms.

Because AI systems are subject to model drift and performance degradation over time, continuous monitoring and periodic auditing are essential.[34]

Scheduled evaluations not only detect algorithmic bias or unintended discrimination but also ensure ongoing compliance with data-protection, anti-discrimination, and sustainability obligations. Transparent reporting is equally crucial: detailed disclosures about AI's role in achieving ESG performance metrics enhance stakeholder trust and facilitate institutional accountability. Including both achievements and identified challenges in annual sustainability reports contributes to a culture of openness and responsible innovation.[35]

Moreover, stakeholder engagement, including dialogue with civil-society organisations, consumer advocates, academic experts, and regulators, broadens

[34] Sandra Wachter, Brent Mittelstadt, & Chris Russell, «Why Fairness Cannot Be Automated…», op. cit.

[35] World Economic Forum, Global Technology Governance Report 2021: Harnessing Fourth Industrial Revolution Technologies in a COVID-19 World, Geneva, WEF, 2021, https://www.weforum.org/reports/global-technology-governance-report-2021 [consult. at 17/07/2025].

the understanding of AI's societal impacts[36] and ensures legitimacy in corporate governance processes. Active participation in international collaboration and knowledge-sharing initiatives, such as the OECD's Global Partnership on Artificial Intelligence (GPAI), promotes the exchange of best practices and helps mitigate transnational risks associated with algorithmic decision-making.[37] In this way, AI governance becomes an integral component of corporate sustainability, reinforcing the alignment between technological innovation, regulatory compliance, and ethical business conduct.

6. The role of stakeholders

6.1. *Policymakers and regulators*

Policymakers must balance the imperative to foster technological innovation with the duty to safeguard fundamental rights and democratic values.

International initiatives most notably the United Nations General Assembly Resolution on "Safe, Secure and Trustworthy AI" (A/78/L.49, 2024) and the OECD Recommendation on Artificial Intelligence (2019), reflect a global movement toward convergence in AI ethics and governance. Yet persistent divergences among national priorities, such as data-privacy protection in the European Union, innovation-driven growth in the United States, or state-centred digital oversight in China, illustrate the difficulty of achieving comprehensive harmonisation. These variations underscore the need for ongoing multilateral coordination to ensure that AI's economic and social benefits are equitably and responsibly distributed across jurisdictions.[38]

For companies, responsible AI adoption requires more than technical compliance; it demands strategic investments in staff training, internal ethics committees, and continuous compliance programmes. The growing prominence of ESG-oriented investment has intensified scrutiny of how AI supports, or potentially conflicts with corporate sustainability strategies. Boards of directors are increasingly expected to demonstrate algorithmic accountability and transparency in sustainability reporting, disclosing both achievements and residual risks. By integrating AI oversight into existing risk-management and corporate-gov-

[36] Future of Life Institute, AI Policy – Global Landscape, last modified 2023, https://futureoflife.org/ai-policy/ [consult. at 17/07/2025].

[37] OECD, State of Implementation of the OECD AI Principles, op. cit.

[38] https://www.un.org/en/summit-of-the-future/global-digital-compact [consult. at 17/07/2025], https://www.un.org/sites/un2.un.org/files/sotf-pact_for_the_future_adopted.pdf [consult. at 17/07/2025].

ernance structures, companies can maintain stakeholder trust and uphold their social licence to operate.[39]

From an operational perspective, embedding AI governance mechanisms across all organisational levels—through policies, standard-operating procedures, and internal auditing—ensures alignment with both regulatory mandates and long-term sustainability goals. Legal practitioners play a pivotal role in this process: they design the contracts, privacy policies, and dispute-resolution mechanisms that govern AI deployment, ensuring compliance with instruments such as the EU AI Act and the Corporate Sustainability Due Diligence Directive. Civil-society organisations, in turn, function as independent watchdogs and advocates, providing external scrutiny, exposing potential abuses, and ensuring that algorithmic systems serve the public interest. Their engagement reinforces corporate accountability and helps embed ethical reflection within the broader ecosystem of sustainable governance.

7. Conclusion

The integration of artificial intelligence into sustainability initiatives aligned with ESG criteria and aimed at fulfilling the SDGs presents both transformative opportunities and significant regulatory and ethical challenges. On one hand, AI enables more efficient resource use, data-driven decision-making, and innovation in climate action and social inclusion. On the other hand, it raises questions concerning bias, privacy, accountability, and human oversight, requiring coherent legal and ethical framework.

By adhering to emerging regulations most prominently the European Union's AI Act and international standards articulated by the United Nations, the OECD, and UNESCO, companies can align AI-driven strategies with global sustainability objectives in a transparent and responsible manner. Policymakers play a crucial role in developing adaptive regulatory regimes that balance protection of fundamental rights with incentives for innovation. Legal practitioners and civil-society organisations complement these efforts by refining governance mechanisms, promoting due diligence, and ensuring effective external accountability.

Concrete measures, such as the creation of interdisciplinary AI task forces, the implementation of continuous algorithmic audits, and the inclusion of AI-specific disclosures in sustainability reports, translate these normative commitments into practice. As AI technologies evolve, the imperative to harmonise innovation with sustainability and human-rights principles will only intensify.

---

[39] Robert G. Eccles, Ioannis Ioannou, & George Serafeim, «The Impact of Corporate Sustainability on Organizational Processes and Performance», op. cit.

Achieving this balance demands close collaboration among policymakers, corporate leaders, regulators, and civil-society actors. By pursuing this collective vision, AI can become not merely a technological instrument but a cornerstone of sustainable development, advancing a more equitable, transparent, and resilient global economy.

O PRINCÍPIO DA LEGALIDADE, A TIPICIDADE FISCAL E A CONFORMAÇÃO DE UM PODER DISCRICIONÁRIO DA ADMINISTRAÇÃO TRIBUTÁRIA EM PORTUGAL: LIMITES FUNCIONAIS À INTELIGÊNCIA ARTIFICIAL

*Luís Manuel Pica**

* Doutor em Ciências Jurídicas Públicas pela Escola de Direito da Universidade do Minho; Professor Adjunto no 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. E-mail: luispica280@gmail.com.

Resumo: O princípio da legalidade fiscal assegura que apenas a Assembleia da República, ou o Governo, mediante autorização legislativa, pode criar impostos, definir os seus elementos essenciais e estabelecer o quadro normativo de atuação da Administração Tributária. Neste contexto, a legalidade não exclui a existência de discricionariedade administrativa, antes a enquadra, permitindo à Administração atuar dentro dos limites traçados pela lei na concretização do interesse público. A progressiva digitalização da atividade tributária e a incorporação de sistemas de inteligência artificial introduzem novos desafios a este equilíbrio. Embora contribuam para a eficiência e para o reforço do controlo fiscal, tais instrumentos podem afetar a transparência e a capacidade de fundamentação das decisões, em virtude da opacidade dos algoritmos e do potencial surgimento de enviesamentos. A aplicação de atos discricionários exige, por isso, um juízo valorativo e interpretativo que os sistemas automatizados dificilmente reproduzem, impondo uma reflexão sobre os limites e garantias da legalidade no contexto da Administração Tributária digital.

Palavras-chave: Legalidade administrativa; Discricionariedade administrativa; Administração Tributária; Inteligência artificial; Limites funcionais.

Abstract: The principle of fiscal legality ensures that only the Assembly of the Republic, or the Government by legislative authorisation, can create taxes, define their essential elements and establish the regulatory framework for the Tax Administration's actions. In this context, legality does not exclude the existence of administrative discretion, but rather frames it, allowing the Administration to act within the limits set by law in the pursuit of the public interest. The progressive digitisation of tax activity and the incorporation of artificial intelligence systems introduce new challenges to this balance. Although they contribute to efficiency and the strengthening of tax control, such instruments can affect transparency and the ability to justify decisions, due to the opacity of algorithms and the potential emergence of biases. The application of discretionary acts therefore requires a value judgement and interpretation that automated systems can hardly reproduce, prompting reflection on the limits and guarantees of legality in the context of digital Tax Administration.

Keywords: Administrative legality; Administrative discretion; Tax Administration; Artificial intelligence; Functional limits.

1. Introdução

O princípio da legalidade, enquanto pedra angular do Estado de Direito Democrático, assume uma relevância ímpar no intrincado domínio do Direito Fiscal. Em Portugal, a sua consagração normativa é cristalina, patente no artigo 103.º, n.º 2, da Constituição **da República Portuguesa**. Enquanto valor axiológico e matricial de um Estado de Direito, este princípio não se limita a afirmar que "os impostos são criados por lei". A sua amplitude exige que o facto tributário deva ser especificado pela lei, a qual deve prever a "incidência [objetiva, subjetiva, temporal e especial], a taxa [nominal], os benefícios fiscais e as garantias dos contribuintes". Esta formulação não é meramente declaratória; ela impõe uma exigência de densidade normativa e uma reserva de lei para a criação e a regulação de qualquer imposto no nosso ordenamento jurídico. A sua génese reside na histórica necessidade de proteger os cidadãos contra a arbitrariedade do poder público, garantindo que a imposição de encargos económicos destinados ao financiamento dos gastos públicos assenta sempre na vontade popular, expressa através dos seus representantes democraticamente eleitos.

A materialização mais evidente do princípio da legalidade fiscal manifesta-se através do princípio da tipicidade. Este último visa assegurar que os elementos essenciais do imposto sejam definidos na lei de forma clara, inequívoca e suficientemente precisa, de modo a afastar qualquer margem para atuações arbitrárias ou casuísticas por parte da Administração Tributária.

Contudo, o cenário económico contemporâneo, caracterizado pela sua crescente complexidade, dinamismo e globalização, aliada à pluralidade de situações com inegável relevância fiscal, torna praticamente inviável uma regulação legal que seja exaustiva e absolutamente fechada. É precisamente neste ponto que surge a necessidade de repensar a tipicidade fiscal, concebendo-a numa dimensão mais "aberta" ou "flexível". Esta abertura admite a utilização de conceitos jurídicos tendencialmente abertos e, concomitantemente, a possibilidade de uma atuação com certo grau de discricionariedade da Administração, sempre, é claro, dentro de limites funcionais e jurídicos rigorosamente definidos. Paralelamente, e de forma inovadora, os limites intrínsecos e funcionais à utilização da inteligência artificial na aplicação de normas tributárias constituem um desafio que se apresenta cada vez mais premente no contexto da digitalização da Administração Pública, designadamente quando utilizada nesta dimensão de cerca discricionariedade administrativa.

2. O princípio da legalidade e a tipicidade como sua concretização

2.1. *Enquadramento*

A natureza reforçada do princípio da legalidade fiscal no ordenamento jurídico português não é uma mera formalidade: *é um pilar da segurança jurídica e da proteção do contribuinte*. Não só a criação de impostos depende de uma lei formal da Assembleia da República – ou Decreto-Lei autorizado do Governo –, como, de forma crucial, os seus elementos essenciais devem ser densamente previstos nessa mesma lei. Embora se apresente como um princípio jurídico transversal aos vários ramos do Direito, o princípio da legalidade pode ser aqui apresentado segundo uma dimensão principiológica material, em que se desenvolve o (sub)princípio da tipicidade fiscal[1].

Esta reserva de lei tem sido consistentemente interpretada de forma estrita pelo Tribunal Constitucional. Um exemplo paradigmático desta interpretação rigorosa encontra-se no Acórdão n.º 128/2009, onde o Tribunal Constitucional afirmou categoricamente que "a densidade normativa dos elementos essenciais da obrigação tributária não pode ser transferida para regulamentos ou outras fontes subalternas". Esta decisão sublinha a ideia de que a sede própria e exclusiva da determinação dos elementos essenciais do imposto é a lei parlamentar – ou o Decreto-Lei autorizado do Governo –, afastando qualquer tentativa de deslegalização da matéria fiscal para diplomas de hierarquia inferior.

Em linha com a consagração constitucional, a Lei Geral Tributária, no seu artigo 8.º, n.º 1, reforça esta exigência ao estabelecer que "a criação de impostos e a definição da respetiva incidência, taxa, liquidação e garantias dos contribuintes só podem ser feitas por lei". Esta disposição espelha inequivocamente a opção do legislador português por um modelo de legalidade estrita, consagrando uma verdadeira reserva absoluta de lei em matéria fiscal. O objetivo é evitar a discricionariedade excessiva e garantir que as decisões que afetam diretamente a esfera patrimonial dos cidadãos sejam tomadas com a maior legitimidade democrática possível.

Dentro deste rigoroso quadro de exigência legal, o princípio da tipicidade fiscal surge como o corolário material da legalidade. Ele implica que os elementos constitutivos do facto tributário – ou seja, os pressupostos que desencadeiam

[1] Sérgio Vasques, *Manual de Direito Fiscal*, 2.ª ed., Almedina, Coimbra, 2019, p. 287; Cfr. ainda Ana Paula Dourado, *O princípio da legalidade fiscal: tipicidade, conceitos jurídicos indeterminados e margem de livre apreciação*, Almedina, Coimbra, 2007, pp. 319 e ss.; *Idem*, «O Princípio da Legalidade Fiscal na Constituição Portuguesa», *Ciência e Técnica Fiscal*, n.º 379, Boletim da Direcção-Geral das Contribuições e Impostos, Lisboa, julho/setembro 1995, pp. 47-97; Alberto Pinheiro Xavier, *Conceito e Natureza do Acto Tributário*, Almedina, Coimbra, 1972, p. 342.

a obrigação de pagar imposto – estejam suficientemente determinados na lei. Esta determinação deve ser tal que permita a sua subsunção automática e previsível aos casos concretos, sem margem para incertezas ou arbitrariedades. Esta exigência é sistematizada na génese de que a tipicidade visa garantir a previsibilidade da atuação fiscal e a proteção dos direitos dos contribuintes. A previsibilidade é, aqui, um valor fundamental, pois permite que o contribuinte saiba, de antemão, quais os seus deveres fiscais e quais as consequências das suas ações, promovendo a segurança jurídica.

Contudo, a tipicidade não ignora que uma abordagem excessivamente rígida e fechada pode tornar-se disfuncional num contexto económico e social em constante mutação. A realidade não se encaixa sempre em categorias pré-definidas com rigidez absoluta. Esta perspetiva é partilhada em outros sistemas jurídicos europeus. Em Espanha, Fernando Sainz Moreno argumenta que "a legalidade fiscal não é um dogma de rigidez, mas um mecanismo de garantia dentro de parâmetros de razoabilidade"[2]. Esta visão aponta para uma legalidade que, embora garantística, não se torna um entrave à eficácia e à justiça tributária. E é precisamente ao enunciado princípio da legalidade fiscal que tende a associar-se uma ideia de tipicidade "aberta" e/ou "fechada", correspondendo à exigência de que a lei fiscal deve estar suficientemente concretizada na determinação do substrato factual tido como relevante para efeitos das normas fiscais. Ou seja, refere-se à capacidade de delimitação do facto tributário e à sua densidade e completude normativa, à atuação da própria administração como forma a encerrar em si os necessários critérios decisórios de aplicação ao caso concreto[3]. Deste modo, quanto mais completa e densa seja a conceção da norma jurídica e do facto tributário que em si encerra, mais fechado será o princípio da tipicidade fiscal.

Embora se compreenda a necessidade de densificação normativa do facto tributário para que, de modo claro e inequívoco, sejam cumpridos os pressupostos legais que se tornam precisos na conformação da realidade objeto de tributação, também é unânime que este princípio da tipicidade deve conceber-se segundo algum grau e elasticidade e de flexibilização. Reclama-se um recorte normativo de alguma "abertura ou elasticidade", no sentido de permitir ao substrato normativo a suficiente abstração e generalidade, para compreender realidades que, dificilmente, seriam concebíveis segundo uma tipicidade fechada[4].

---

[2] Fernando Sainz de Moreno, *Conceptos jurídicos, interpretación y discrecionalidad administrativa*, Civitas, Madrid, 1976, p. 312.

[3] Alberto Pinheiro Xavier, *Conceito e Natureza do Acto Tributário*, Almedina, Coimbra, 1972, pp. 342 e ss.; Ana Paula Dourado, *O princípio da legalidade fiscal…*, *op. cit.*, pp. 103 e ss.; José Casalta Nabais, *O Dever Fundamental de Pagar Impostos*, Almedina, Coimbra, 2012, pp. 368 e ss.

[4] Ana Paula Dourado, *O princípio da legalidade fiscal…*, *op. cit*.

Apesar dos condicionalismos jurídico-constitucionais que derivam do princípio da tipicidade em matéria tributária (cfr. artigo 103.º, n.ºs 2 e 3, da Constituição da República Portuguesa), a lei fiscal deve conter em si mesma uma valoração definitiva das realidades sobre as quais versa, excluindo elementos a ela estranhos e que possam conhecer vontades externas e alheias à vontade do órgão legislativo primário, enquanto detentor do poder de criar impostos. Exige-se, assim, uma densificação suficiente e completa da descrição dos pressupostos da tributação, de modo que a norma seja satisfatoriamente densa, determinável e abrangente, que permita a produção de efeitos jurídicos vinculativos para o seu destinatário[5].

Ora, num contexto de sociedade altamente desenvolvida e cuja evolução é vertiginosa, as preocupações que subjazem à necessidade de densificação normativa implicam o seu confronto com a dificuldade de dar resposta a todas as situações e factualidades económicas que se vão desenvolvendo e complexificando. Ou seja, um conceito normativo extremadamente fechado contende fortemente com a função primordial e imediata que será a ordenação da esfera das relações entre o poder legislador e executivo, vinculando a administração *inteligente* segundo um conceito restrito que frustre os objetivos da tributação. Daí que se reclame por uma abertura e flexibilização que coloque em concordância com uma busca adequada de equilíbrio e tipicidade, recortada com alguma plasticidade semântica que compreenda a capacidade de abarcar realidades que se expressem no conceito e que se assumam como integrantes na capacidade contributiva[6].

Assim, a tipicidade fiscal contemporânea não é encarada como um conceito absolutamente estático e fechado, mas sim como um modelo dinâmico que deve procurar um equilíbrio virtuoso entre a densidade normativa e a flexibilidade interpretativa. Este equilíbrio é crucial para que o sistema fiscal possa, por um lado, oferecer segurança e previsibilidade aos contribuintes e, por outro, ser eficaz na tributação de realidades económicas complexas e em constante evolução, sem se tornar obsoleto ou injusto. A capacidade de lidar com a inovação e as novas formas de riqueza exige uma tipicidade que, embora salvaguarde os princípios essenciais, não seja um obstáculo à adaptação e à equidade.

---

[5] Alberto Pinheiro Xavier, *Conceito e Natureza do Acto Tributário*, *op. cit.*, pp. 342 e ss.; Ana Paula Dourado, *O princípio da legalidade fiscal…*, *op. cit*., pp. 103 e ss.

[6] Ana Paula Dourado, «O Princípio da Legalidade Fiscal na Constituição Portuguesa», *Ciência e Técnica Fiscal*, n.º 379, Boletim da Direcção-Geral das Contribuições e Impostos, Lisboa, julho/setembro de 1995, p. 84.

2.2. *Conceitos tendencialmente abertos e poder discricionário de uma Administração Tributária "artificialmente inteligente"*

Como se deixou referido anteriormente, a inevitável admissão de conceitos tendencialmente abertos na legislação fiscal é um corolário da complexidade inerente à vida económica e, simultaneamente, um reconhecimento da dificuldade (e até da indesejabilidade) de o legislador prever exaustivamente todas as situações passíveis de tributação. Termos como "rendimento", "atividade profissional", "transmissão onerosa", "gastos" ou "prejuízo e injustiça grave" não são meras formulações retóricas. Os conceitos tendencialmente abertos, genéricos e abstratos com alguma elasticidade abrem, de facto, um espaço crucial para a atuação valorativa e interpretativa da Administração Tributária. Estas expressões exigem uma interpretação contextualizada e uma apreciação casuística por parte dos serviços fiscais, que, ao aplicá-las, têm de ir além da mera subsunção formal. Esta realidade tem sido amplamente discutida, salientando a "elasticidade semântica" destes conceitos, os quais permitem uma maior adaptabilidade da norma à realidade, mas exige uma fundamentação acrescida por parte da Administração Tributária.

Para além disto, a Administração é, atualmente, legitimada a atuar com certo grau de discricionariedade no âmbito das suas funções. Não se dirá que é uma legitimação total, mas sim uma prerrogativa sectorialmente localizada pelo legislador que subjetiviza a atuação da Administração. Por isso, este poder discricionário da Administração não é, nem pode ser, ilimitado. A jurisprudência tem sido consistentemente clara ao afirmar que a discricionariedade, mesmo quando prevista na lei, só pode ser exercida dentro dos limites impostos pela própria lei e segundo os princípios gerais do Direito Administrativo, nomeadamente os da proporcionalidade, razoabilidade, boa administração, imparcialidade e igualdade. Um exemplo eloquente encontra-se no Acórdão do Tribunal Central Administrativo do Norte de 02.07.2021, onde se reafirma que o exercício do poder discricionário "não confere à Administração um poder arbitrário, antes a vincula aos fins que a lei visa prosseguir, e a sujeita aos princípios gerais do direito, designadamente aos princípios da proporcionalidade e da boa-fé"[7]. Este controlo jurisdicional da discricionariedade é a garantia última para o contribuinte contra atuações abusivas.

A própria Lei Geral Tributária consagra diversas situações onde a Administração atua com uma margem de apreciação valorativa.

Por exemplo:

---

[7] Acórdão do Tribunal Central Administrativo do Norte de 02.07.2021, proc. 00718/17.9BECBR, Relator: Desembargadora Helena Ribeiro, disponível em http://www.dgsi.pt [consultado em 04.08.2025].

1) Dispõe o artigo 38.º, n.º 2, da Lei Geral Tributária que a Administração está habilitada a desconsiderar atos ou negócios jurídicos que, embora formalmente válidos, visem essencialmente a obtenção de uma vantagem fiscal que, em face das circunstâncias, não seria concedida. Esta qualificação de atos ou negócios como simulados ou artificiais implica uma profunda análise das intenções das partes e dos efeitos económicos reais subjacentes, indo para além da mera forma jurídica. O Tribunal de Justiça da União Europeia, em casos como o conhecido Halifax (processo C-255/02)[8], tem vindo a desenvolver o conceito de "abuso de direito" em matéria de IVA, que se assemelha a esta norma anti-abuso, exigindo uma análise objetiva das operações e uma prova da intenção de obter uma vantagem fiscal indevida.

2) Por outro lado, o artigo 87.º da Lei Geral Tributária dispõe que a utilização de métodos indiretos de avaliação da matéria coletável é um dos exemplos mais claros do poder discricionário da Administração Tributária, especialmente quando a contabilidade do contribuinte se revela omissa, deficiente ou manifestamente inverídica. Embora a lei estabeleça critérios para a sua aplicação (como a impossibilidade de determinar o rendimento ou valor patrimonial por métodos diretos), a fixação concreta do valor tributável segundo métodos indiretos envolve uma avaliação estimativa, que deve ser fundamentada em indicadores objetivos e passível de controlo jurisdicional, bem como a possibilidade de estabelecimento de um acordo bilateral entre a administração e o contribuinte.

Em perspetiva, a admissão de conceitos abertos, bem como a legitimação de certa discricionariedade administrativa são práticas comuns em sistemas fiscais desenvolvidos, podendo mesmo dizer-se que estamos perante uma vinculação tendencialmente neutra, pois a Administração age vinculada à lei, mas com certa discricionariedade na medida em que não atue contrariamente ao que esteja delimitado normativamente. A Administração Tributária não deixa de ser um órgão da Administração Pública com competências específicas no âmbito da cobrança, fiscalização e aplicação das normas tributárias, tendo por base o princípio da legalidade administrativa (artigo 266.º, n.º 2, da Constituição da República Portuguesa). A sua atuação está sujeita aos limites constitucionais que impõem vinculação à lei, ao interesse público e à prossecução dos direitos e garantias dos contribuintes.

A própria Constituição estrutura o poder do Estado segundo o princípio da separação e interdependência dos poderes (artigo 111.º da Constituição da República Portuguesa), reservando ao poder legislativo a criação de tributos e ao executivo a sua aplicação. A Administração Tributária, como entidade admi-

[8] Acórdão do Tribunal de Justiça das Comunidades Europeias, de 21 de Fevereiro de 2006. Processo C – 225/02 – Halifax, disponível em https://curia.europa.eu/juris/document/document.jsf;jsessionid=BECABF6D142185C3B7E6608B4576519F?text=&docid=56198&pageIndex=0&doclang=pt&mode=lst&dir=&occ=first&part=1&cid=7394834 [consultado em 04.08.2025].

nistrativa inserida na estrutura do Ministério das Finanças, atua apenas dentro do quadro normativo previamente estabelecido pela lei tributária, sem margem para inovar ou criar obrigações ex novo[9]. Ou seja, a Administração não pode, sob pretexto de interpretar ou aplicar a lei, subverter a vontade do legislador ou extravasar os limites por ele traçados. Como refere Domingos Pereira de Sousa[10], "a discricionariedade é, definitivamente, um dos temas mais controversos na doutrina do direito público, em particular quando se analisam a sua localização e a forma do seu exercício no Estado de Direito contemporâneo assente na denominada separação de poderes, entendida hoje como divisão de funções derivadas do exercício do poder político intrínseco à soberania popular, nos termos constitucionais".

Deste modo, o interesse público constitui o guia norteador de toda a atividade administrativa, pautando a mesma segundo toda a principiologia inerente a todo o bloco normativo (cfr. artigo 3.º, n.º 1, do Código do Procedimento Administrativo). Portanto, a atuação da Administração Pública orienta-se segundo uma égide de realização e salvaguarda pelo interesse público, enquanto matriz central de toda a atuação administrativa[11], e, para o efeito, essa égide matricial exige o respeito pelos direitos e interesses legalmente protegidos dos cidadãos (cfr. artigo 266.º, n.º 1, da Constituição da República Portuguesa). Assim, a atuação da Administração não pode deixar de respeitar os limites normativamente impostos, ainda que legitimem uma atuação balizada segundo um certo grau de discricionariedade, exigindo-se uma atuação normativamente vinculada e pautada pela legalidade administrativa.

2.3. *Discricionariedade administrativa* ***vs.*** *vinculação legal*

A atuação da Autoridade Tributária e Aduaneira, ainda que, em certos domínios, o legislador lhe reconheça uma margem de livre apreciação – como sucede, *inter alia*, na avaliação indireta da matéria tributável (artigos 87.º e ss. da Lei Geral Tributária) ou na qualificação de negócios simulados (artigo 38.º, n.º 2, da Lei Geral Tributária) –, permanece integralmente subordinada à ordem jurídica, não podendo ser confundida discricionariedade com liberdade arbitrária.

---

[9] Joaquim Freitas da Rocha; Hugo Flores da Silva, Teoria Geral da Relação Jurídica Tributária, Almedina, Coimbra, 2021, pp. 41 e ss.

[10] Domingos Pereira de Sousa, «Legalidade, Vinculação e Discricionariedade na Administração Tributária», De Legibus – Revista de Direito da Universidade Lusófona, n.º 28, Jan. de 2021, p. 34.

[11] Diogo Freitas do Amaral, Curso de Direito Administrativo, Vol. II, Almedina, Coimbra, 2016, p. 38.

Como sustenta Jorge Miranda[12], "a discricionariedade não se confunde com arbítrio: há sempre uma moldura legal, uma vinculação mínima que não pode ser ultrapassada". Quer isto dizer que o espaço de liberdade conferido pela lei à Administração não significa ausência de limites, mas sim uma liberdade juridicamente delimitada, balizada por princípios e regras superiores.

A vinculação da Autoridade Tributária e Aduaneira decorre, desde logo, do princípio constitucional da legalidade (artigos 103.º e 266.º, n.º 2, da Constituição da República Portuguesa), do dever de prosseguir o interesse público com respeito pelos direitos e interesses legalmente protegidos dos cidadãos (artigo 266.º, n.º 1, da Constituição da República Portuguesa) e da obrigação de fundamentação (artigo 77.º, n.º 1, da Lei Geral Tributária), o que implica que a margem de apreciação se subordine aos princípios da proporcionalidade, razoabilidade, justiça e imparcialidade (artigo 55.º da Lei Geral Tributária).

Exige-se que a margem de apreciação que a Autoridade Tributária e Aduaneira deve exercer seja concretizada segundo juízos de proporcionalidade, razoabilidade e fundamentação das suas decisões, sob pena de invalidade do ato tributário, não constituindo a discricionariedade administrativa um espaço de indiferença jurídica, mas sim uma liberdade delimitada por parâmetros legais e constitucionais, sindicável pelos tribunais. A discricionariedade é, assim, sindicável, tanto no plano do cumprimento dos pressupostos legais, como na aferição da razoabilidade e proporcionalidade da decisão administrativa, sempre que haja indícios de desvio de poder ou erro manifesto. Não se trata de uma sindicância sobre o mérito ou conveniência, nem, tampouco, uma substituição do órgão administrativo, mas sim uma fiscalização dos parâmetros normativos que balizam a atuação administrativa[13].

Assim, não será despiciendo afirmar que o poder discricionário da Autoridade Tributária e Aduaneira é um poder cuja génese reside no bloco legal e normativo que lhe impõe as condutas admissíveis e possíveis, porquanto resulta de uma autorização expressa que o legislador lhe atribui para concretização dos desideratos matriciais de um Estado de Direito. A concessão de uma Administração capaz de atuar perante um espetro discricionário devidamente delimitado através da intenção de lhe conferir uma margem de discricionariedade, através de diferentes mecanismos normativos dados através de normas habilitantes, é aquilo que se pode designar de "canais de abertura discricionária".

Assim, a norma habilitante atribui à Administração um espaço próprio de apreciação, permitindo-lhe formular juízos autónomos acerca dos factos re-

[12] Jorge Miranda, *Manual de Direito Constitucional*, Vol. IV, 6.ª ed., Coimbra Editora, Coimbra, 2021, p. 235.

[13] Acórdão do Supremo Tribunal de Justiça de 24.02.2022, proc. 39/20.0YFLSB, Relatora: Conselheira Maria Olinda Garcia, disponível em http://www.dgsi.pt [consultado em 04.08.2025].

levantes e da qualificação de pessoas, situações ou circunstâncias concretas. Em qualquer das hipóteses, a decisão assenta na existência de uma disposição legal que, por meio de uma cláusula de abertura discricionária, confere à Administração competência para ponderar os elementos específicos do caso e, a partir dessa ponderação, definir o conteúdo material e a solução final mais apropriada.

Apesar da margem de liberdade conferida pela discricionariedade administrativa, esta encontra sempre *limites internos*, designadamente quanto à competência do órgão que decide e ao fim de interesse público que deve orientar a atuação — elementos que permanecem juridicamente vinculados. No quadro de uma *norma permissiva*, a Administração dispõe da faculdade de agir ou de se abster, configurando-se a chamada *discricionariedade de ação ou de decisão*. Nesta hipótese, não há um dever jurídico de intervir, mas sim a possibilidade de o fazer, guiada por critérios de oportunidade e conveniência; já no âmbito de uma *norma optativa*, o legislador define previamente um conjunto de medidas possíveis, cabendo à Administração escolher, entre essas opções, a que melhor se ajusta às particularidades do caso concreto. Além disso, a própria norma pode estabelecer um intervalo temporal para a atuação administrativa, concedendo liberdade quanto à determinação do momento exato da prática do ato.

Mas esta legitimação de uma discricionariedade administrativa encontra, também, *limites externos* ligados à tutela de valores fundamentais inseridos num Estado de Direito. A proteção dos direitos fundamentais dos contribuintes constitui, neste contexto, um limite material intransponível, abrangendo o princípio da igualdade (artigo 13.º da Constituição da República Portuguesa), o direito de propriedade (artigo 62.º da Constituição da República Portuguesa) e a tutela jurisdicional efetiva (artigo 20.º da Constituição da República Portuguesa), com projeção internacional através da Convenção Europeia dos Direitos Humanos.

A compreensão dos direitos fundamentais cujo reflexo também deve ser feito em matéria tributária só é legítima se obedecer ao princípio da proporcionalidade em todas as suas vertentes — adequação, necessidade e proporcionalidade em sentido estrito — e respeitar a proteção da confiança dos contribuintes (artigos 2.º e 18.º da Constituição da República Portuguesa). Assim, qualquer atuação da Autoridade Tributária e Aduaneira que, a coberto de uma suposta discricionariedade, imponha medidas arbitrárias, desproporcionais ou destituídas de fundamentação suficiente viola os princípios estruturantes do Estado de Direito fiscal e está sujeita a anulação pelos tribunais administrativos ou arbitrais, no quadro de uma tutela jurisdicional efetiva que constitui garantia essencial do contribuinte.

Acrescendo a isto, a inclusão da inteligência artificial na Administração Pública, incluindo na Administração Tributária, levanta questões fundamentais sobre a compatibilidade entre a automação da decisão administrativa e os prin-

cípios estruturantes do Direito Fiscal, nomeadamente o princípio da legalidade, da tipicidade fiscal e da capacidade contributiva.

3. A digitalização, a inteligência artificial e os meios eletrónicos na atividade Administrativa Tributária

3.1. *Enquadramento*

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

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 desenvolver-se 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 de 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.

Neste contexto de globalização, crise e transformação tecnológica, a Administração Pública vê-se confrontada com uma crescente necessidade de adaptação e resposta célere a novas realidades sociais e económicas. Essa exigência acentua a importância da discricionariedade administrativa enquanto poder de escolha legítimo dentro dos limites da lei, permitindo à Administração concre-

tizar o interesse público face a situações que o legislador não consegue antecipar integralmente.

A transição para uma Administração eletrónica não elimina esse espaço decisório, mas antes o reconfigura, colocando novos desafios quanto à forma de exercer a discricionariedade num ambiente dominado por procedimentos automatizados e decisões baseadas em dados. Importa, por isso, refletir sobre como assegurar que o exercício da discricionariedade administrativa, ainda que mediado por tecnologias, continue a respeitar os princípios da legalidade, proporcionalidade e fundamentação que estruturam o Estado de Direito administrativo.

Por essa razão, discute-se, no âmbito da atividade administrativa e da organização da Administração Pública, a ideia de uma *Administração eletrónica*[14]. É que o advento das novas tecnologias deu origem a uma revolução tecnológica, com inegáveis consequências no domínio económico e social que estão 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 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.º 1, da Lei Geral Tributária já prevê a utilização das novas tecnologias no procedimento tributário.

Com efeito, a Administração Pública eletrónica não é mais do que uma combinação entre Administração Pública e os sistemas de tecnologia de informação que a integram, 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[15]. Do mesmo modo, o procedimento administrativo eletrónico não é diferente do procedimento administrativo comum, mas sim o resultado da aplicação das tecnologias de informação e comunicação

---

[14] Miguel Prata Roque, «O Nascimento da Administração Eletrónica num Espaço Transnacional (Breves notas a propósito do projeto de revisão do Código do Procedimento Administrativo)», *e-Pública*, Vol. I, n.º 1, janeiro de 2014, pp. 307-321.

[15] Santi Romano, *Principii di Diritto Amministrativo Italiano*, CEDAM, Milão, 1937, p. 77.

a este último, com as novas possibilidades de melhor cumprimento do princípio da eficiência que daí advêm[16].

Vejamos, com maior pormenor, a sua inserção no domínio da aplicação das normas administrativas e tributárias.

3.2. *A atividade administrativa e a digitalização, a inteligência artificial e os meios eletrónicos*

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 análogas, sendo, por vezes, utilizados indistintamente. Por exemplo, se tomarmos como exemplo o disposto no n.º 1 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.

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:

i) 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.

ii) Por outro lado, a *automatização* traz um *plus*, pois existe uma transformação mais profunda, introduzindo sistemas de apoio à decisão, como a inteligência artificial, que podem substituir os agentes humanos na tomada de decisões e na execução dos atos procedimentais. Aqui, a inteligência artificial vai além da simples digitalização, utilizando algoritmos baseados em critérios pré-estabelecidos para interpretar e aplicar as normas jurídicas, formulando pro-

[16] Marcelo Rebelo de Sousa; André Salgado de Matos, *Direito Administrativo Geral – Tomo I – Direito Administrativo Geral. Introdução e Princípios Fundamentais*, 3.ª ed., Dom Quixote, Lisboa, 2008, pp. 38 e ss.

jetos de decisões, ideias ou conclusões que podem ser adotadas no procedimento[17]. 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: (i) 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 (incluindo os algoritmos) farão o que os burocratas faziam; (ii) a realidade na qual a modernização, a automatização e a integração de 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"[18]; (iii) 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.

Embora seja este o nosso entendimento, não podemos deixar de concluir pela ideia que a *digitalização* do procedimento traduz-se num mero ato de integração de sistemas tecnológicos destinados a digitalizar o suporte dos atos praticados no procedimento tributário, sem que exista, conquanto, um suporte na atividade decisória; pelo contrário, a adoção de meios de inteligência artificial,

---

[17] Neste sentido, veja-se a definição dada pelo artigo 3.º, n.º 1, do Regulamento 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".

[18] Uma posição mais restritiva pode ser encontrada em Alexandre Dias Pereira, segundo o qual "[d]o que fica dito poderia encontrar-se o «código-fonte» do sistema jurídico na Constituição. Ao intérprete da lei ordinária estaria reservada a possibilidade de, quando muito, «desvendar» princípios consignados na Constituição, ainda que não explícita ou expressamente consagrados, e mesmo assim sob o controlo do Tribunal Constitucional. Com efeito, princípios consignados serão, por muito fluida que seja a palavra consignar, princípios previstos na Constituição ou que pelo menos (o que já não será assim tão pouco) não ofendam o espírito do sistema constitucional vigente". Para além disso, "[o]s sistemas informáticos de inteligência artificial prometem uma justiça sem arbítrio. Árbitro seria apenas o programador político do sistema que implementaria tecnicamente no software o programa político de domínio e ordenação social contido no código-fonte constitucional e das leis. Juízos só os provenientes da vontade política do legislador. A máquina limitar-se-ia a reproduzi-los mecanicamente nos casos concretos por via de aplicação silogística" – Alexandre Dias Pereira, «*Ius Ex Machina*? Da Informática Jurídica ao Computador-Juiz», *RJLB*, Ano 3, n.º 1, 2017, p. 81.

procurando a *automatização* procedimental, implica integrar um sistema capaz de ler dados que a Administração tem em sua posse, os quais são tratados massivamente para efeitos de produzir uma decisão que servirá de base à prática do ato decisório final.

4. A inteligência artificial e a ausência de discernimento valorativo na gestão do sistema fiscal: limitações da inteligência artificial na aplicação de conceitos indeterminados

Os conceitos jurídicos que atribuem poderes discricionários à Administração Tributária exigem uma contextualização e adaptação ao caso concreto que pressupõem raciocínio interpretativo e valorativo factual, jurídico e social. Tais conceitos são formulados precisamente para permitir uma adequação da norma às múltiplas situações que escapam à previsão exaustiva do legislador[19]. Dogmaticamente, procura-se que a densificação normativa, a sua completude e precisão normativa não possa ser entendida como absoluta e inflexível, outorgando-se alguma flexibilização e abertura que permita dar cumprimento aos demais desideratos constitucionais como a igualdade e a capacidade contributiva[20].

---

[19] José Casalta Nabais, *O Dever Fundamental de Pagar Impostos*, *op. cit.*, pp. 373 e ss.; João Sérgio Ribeiro, *Tributação Presuntiva do Rendimento*, Almedina, Coimbra, 2014, p.123; José Luís Saldanha Sanches, «A segurança jurídica no Estado social de direito», *Ciência e Técnica Fiscal*, n.os 310/312, 1984, pp. 299 e ss.; João Caupers, *Introdução ao Direito Administrativo*, 10.ª ed., Âncora Editora, Lisboa, 2009, pp. 89 e ss.; Marcelo Rebelo De Sousa; André Salgado De Matos, *Direito Administrativo Geral, Introdução e Princípios Fundamentais – Tomo I*, Dom Quixote, Lisboa, 2004, pp.176 e ss.; Diogo Freitas do Amaral, *Curso de Direito Administrativo*, Vol. II, 4.ª ed., Almedina, Coimbra, 2021, pp. 84 e ss.; Mário Aroso de Almeida, *Teoria Geral do Direito Administrativo*, 5.ª ed., Almedina, Coimbra, 2018, pp. 54-55.

[20] "Mas, do outro lado, o princípio da igualdade tributária reclama que os conceitos tenham a abertura ou plasticidade semântica suficiente para poder abarcar as realidades que expressam a capacidade tributária elegida, os níveis de riqueza ou de rendimento tributando, e que esse objectivo possa ser realizado não só no plano abstracto da previsão dos tipos tributários, mas também no plano da sua aplicação concreta, em que se situam o combate à evasão fiscal e a praticabilidade do sistema. Na verdade, sem uma estruturação conceitual apta a colher todas as virtualidades que o facto tributário é susceptível de assumir; sem uma eleição de conceitos que permitam surpreender a existência da riqueza e dos rendimentos e, por fim, sem a construção de um sistema exequível sob o ponto de vista da praticabilidade, não pode haver um efectivo cumprimento do princípio da igualdade tributária. São estes, essencialmente, os aspectos que, segundo a doutrina, justificam o uso dos conceitos jurídicos indeterminados, de «certas cláusulas gerais», de «conceitos tipológicos» (*Typusbegriffe*), de «tipos discricionários» (*Ermessentatbestände*) e de certos conceitos que atribuem à administração uma margem de valoração, os designados «preceitos de poder» (*Kann-Vorschrift*)" – Acórdão do Tribunal Constitucional n.º 127/04, proc. n.º 136/03, Relator: Conselheiro Benjamin Rodrigues, disponível em https://www.tribunalconstitucional.pt/tc/acordaos/20040127.html [consultado em 06.08.2025].

A questão da aplicação de conceitos jurídicos abertos e que permitem uma certa margem discricional de atuação administrativa cuja atividade é ancorada por sistemas de inteligência artificial inscreve-se no cerne da tensão entre a racionalidade jurídico-normativa e a racionalidade tecnológica. Isto porque o problema de automatizar o ato discricionário é que, ao fazê-lo, deixa de existir a opção do agente ou funcionário da administração, isto é, o resultado passa a ser vinculado, tendo em vista que, na automação, o sistema computacional é desenhado para apresentar uma resposta segundo as regras previamente introduzidas, perdendo-se a flexibilidade pretendida nas atuações discricionárias.

Por outro lado, o exercício da discricionariedade implica a necessidade de empatia. Por exemplo, ao avaliar conceitos jurídicos como boa conduta e boa-fé, no exercício da equidade na revisão *ex officio* ou revogação, deve ser levada em consideração a empatia. E o Direito, ao contrário de sistemas normativos fechados, é estruturado como um sistema aberto, no qual a linguagem normativa cumpre simultaneamente funções prescritivas, interpretativas e valorativas. Essa abertura é particularmente visível nos conceitos jurídicos valorativos e de atuação discricionária, os quais representam uma estratégia legislativa essencial para assegurar a adaptabilidade do ordenamento a situações concretas e que exigem uma avaliação valorativa casuística, traduzindo-se numa liberdade de escolha dentro dos limites normativos fixados.

Por isso, uma administração encorada numa estrutura funcional baseada em sistemas de inteligência artificial levanta três grandes limitações na aplicação de conceitos jurídicos indeterminados:

I. *Falta de hermenêutica normativa*

Em primeiro lugar, a ausência de uma verdadeira capacidade hermenêutica por parte da inteligência artificial coloca em evidência uma problemática central: *a da interpretação jurídica.*

A máquina, ainda que dotada de mecanismos sofisticados de *machine learning* ou *deep learning*, atua por meio de correlações estatísticas, identificando padrões a partir de bases de dados e replicando-os em novos contextos. A sua "aprendizagem" é, em rigor, uma otimização probabilística, não uma compreensão semântica. Ou seja, a inteligência artificial não lê nem compreende, apenas calcula probabilidades de associação entre sinais. Assim, mesmo quando aparenta reproduzir resultados que se aproximam de uma decisão humana, trata-se de um processo de simulação estatística, e não de interpretação no sentido jurídico.

Ora, a interpretação do Direito não se esgota na literalidade das palavras nem no reconhecimento de padrões formais. A hermenêutica jurídica implica uma reconstrução do sentido da norma em conformidade com a Constituição, com os princípios gerais do Direito e com o contexto social e histórico em que se

insere. Como defende Gomes Canotilho, interpretar significa recriar normativamente o texto legal à luz da Constituição e da ordem axiológica do sistema jurídico[21], exigindo-se uma compreensão teleológica do Direito constituído, isto é, a interpretação deve atender à finalidade normativa e ao valor que se procura proteger, o que transcende qualquer cálculo probabilístico.

Quando a norma legitima a Administração a atuar segundo critérios decisórios de índole subjetiva ou valorativa, a interpretação exige a articulação entre o texto legal, a situação concreta e os valores constitucionais em jogo. É precisamente aqui que reside a maior dificuldade para uma atuação algorítmica: a inteligência artificial carece da capacidade de ponderar princípios em conflito, de atribuir relevância axiológica a determinados factos e de interpretar a realidade social em termos normativos. A Administração, na sua atuação tradicional, exerce uma função que é simultaneamente cognitiva e valorativa: recolhe os factos, enquadra-os juridicamente e decide segundo critérios de proporcionalidade, justiça e adequação constitucional. Trata-se de um exercício em que a objetividade dos factos nunca é suficiente, sendo indispensável um momento de apreciação subjetiva, na qual a Administração não é uma mera executora da lei, mas verdadeiramente uma intérprete do Direito.

Já uma Administração baseada em atuações puramente objetivistas e mecanizadas – como a que resultaria do uso exclusivo da inteligência artificial – limita-se a reproduzir regularidades formais e a aplicar esquemas previamente definidos, sem captar a dimensão axiológica da norma. O risco traduz-se, assim, numa atuação uniformemente estatística, em detrimento de uma justiça material e de interesse público. Uma decisão correta no plano formal pode revelar-se injusta no plano substancial, por não ter em conta a especificidade do caso e os valores axiológicos que a norma visa proteger.

Neste sentido, a aplicação de conceitos indeterminados por sistemas de inteligência artificial confronta-se com um limite estrutural. Por mais avançado que seja o algoritmo, ele não pode realizar a mediação hermenêutica entre o texto normativo, os princípios constitucionais e o caso concreto. A máquina não interpreta, apenas processa dados. E onde não há interpretação, não há verdadeira aplicação do direito, mas apenas a sua simulação técnica. Por isso, não podemos deixar de referir que a interpretação das normas jurídicas e a leitura dos factos numa Administração legitimada a atuar segundo uma tendência discricionária conceitos é uma função essencialmente hermenêutica que não pode ser delegável a máquinas.

---

[21] J. J. Gomes Canotilho, *Direito Constitucional e Teoria da Constituição*, 7.ª ed., Almedina, Coimbra, pp. 258-259.

II. *Opacidade algorítmica (**black box**)*

Em segundo lugar, a utilização da inteligência artificial numa Administração legitimada para atuar segundo certo grau de discricionariedade suscita problemas particularmente graves de *opacidade algorítmica*, também conhecida como *black box*.

Esta opacidade decorre do facto de muitos sistemas de *machine learning* e, sobretudo, de *deep learning*, funcionarem através de redes neuronais artificiais complexas, cujo processo decisório não é facilmente compreensível nem sequer pelos próprios programadores que os conceberam. O resultado é que, embora a máquina produza uma decisão ou recomendação, os critérios concretos que conduziram a esse resultado permanecem ocultos ou ininteligíveis para terceiros.

Este fenómeno levanta problemas à luz do princípio da fundamentação dos atos administrativos, consagrado no artigo 268.º, n.º 3, da Constituição da República Portuguesa, no artigo 152.º do Código do Procedimento Administrativo e no artigo 77.º da Lei Geral Tributária. A exigência de fundamentação não é uma mera formalidade, constitui uma garantia essencial do contribuinte, permitindo-lhe compreender as razões da decisão, avaliar a sua legalidade e, se for caso disso, contestá-la em sede jurisdicional. A fundamentação constitui, assim, um requisito de transparência e de racionalidade, assegurando que a Administração não atua de forma arbitrária ou obscura. Como se denota, estabelece-se como uma verdadeira garantia dos administrados e do Estado de Direito Democrático, assegurando transparência, imparcialidade e a possibilidade de controlo dos atos administrativos[22].

Todavia, não se deixará de assumir que a fundamentação dos atos tributários emanados através de sistemas de inteligência artificial deverá, ainda, ser uma garantia reforçada. Por isso, a fundamentação dos atos emanados deve permitir uma *explicabilidade interna (ou material)*, a qual diz respeito à fundamentação do ato tributário emanado, mas, também, uma *explicabilidade externa (ou funcional)*, designadamente quanto ao algoritmo utilizado e à forma como este produziu a decisão emanada.

Esta *explicabilidade externa (ou funcional)* acaba por ser consagrada no artigo 14.º, n.º 2, alínea g), do Regulamento Geral sobre a Proteção de Dados, bem como no artigo 86.º, n.º 1, do Regulamento Inteligência Artificial, aprovado pelo Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024. O que se refere neste princípio da explicabilidade exter-

[22] José Carlos Vieira de Andrade, *O Dever da Fundamentação Expressa de Actos Administrativos*, Almedina, Coimbra, 1991, p. 79; Débora Melo Fernandes, «O Princípio da Transparência Administrativa: Mito ou Realidade?», *Revista da Ordem dos Advogados*, Ano 75, jan./jul. de 2015, p. 443.

na é que a Administração Tributária implemente medidas que possibilitem aos contribuintes o acesso às informações sobre a lógica subjacente às decisões automatizadas[23]/[24]. Trata-se, aliás, de evitar aquilo que é designado por *black box* ou *caixa negra*, e que corresponde à impossibilidade de a Administração Tributária justificar o motivo e os fundamentos da decisão emanada pelos sistemas de inteligência artificial por si utilizados nas atividades de gestão fiscal[25].

Pelo contrário, a *explicabilidade interna (ou material)* não deixará de buscar a sua axiologia do Direito Administrativo clássico. É que não podemos esquecer que o procedimento tributário, ainda que pautado por sistemas de inteligência artificial, está subordinado pelas regras e princípios consolidados no ordenamento jurídico português. Os direitos adquiridos aos longos dos anos não podem ser simplesmente eliminados por deslumbramentos de simplicidade ou eficiência das novas tecnologias. Por isso, o direito à fundamentação dos atos administrativos constitui um valor matricial do Estado de Direito que não pode ser suprido ou limitado, e, no seu desenvolvimento, obriga a que as decisões proferidas no âmbito do procedimento tributário permitam conhecer a exposição dos motivos que justificam a resolução tomada (*iter cognitivo*), permitindo ao administrado compreender as razões que levaram a tal decisão e, se necessário, impugná-la[26].

Em sintonia com tal entendimento, o desenvolvimento normativo deste princípio acaba por ser dado, geralmente, pelo artigo 153.º, n.º 1, do Código do Procedimento Administrativo, o qual estabelece que "*a fundamentação deve ser expressa, através de sucinta exposição dos fundamentos de facto e de direito da*

---

[23] Luís Manuel Pica, *A Inteligência Artificial no Direito Tributários: Fundamentos e Limites Constitucionais*, Almedina, Coimbra, 2023, pp. 181 e ss.; Angelo Giuseppe Orofino, «The Implementation of the Transparency Principle in the Development of Electronic Administration», *European Review of Digital Administration & Law – Erdal*, Volume 1, Issue 1-2, June/December 2020, Aracneeditrice, pp. 123-142.

[24] Com muito interesse, veja-se a Sentença do Tribunal de Haia de 05.02.2020, proc. n.º C-09-550982-HA ZA 18-388, disponível em https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878 [consultado em 06.08.2025].

[25] Fernando Serrano Antón, «Fiscalidad y Robótica: funcionalidades disruptivas en el Derecho tributario», in *Fiscalidad e Inteligencia Artificial: Administración Tributaria y Contribuyentes en la era digital*, Thomson Reuters Aranzadi, p. 52; Yavar Bathaee, «The Artificial Intelligence Black Box and the Failure of Intent and Causation», *Harvard Journal of Law & Technology*, Vol. 31, Number 2, Spring 2008, pp. 890-938.

[26] No Acórdão do Supremo Tribunal Administrativo de 19.02.2003, 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" – Proc. n.º 762/02, Relator: Conselheiro António Madureira, disponível em http://www.dgsi.pt [consultado em 06.08.2025].

decisão, podendo consistir em mera declaração de concordância com os fundamentos de anteriores pareceres, informações ou propostas, que constituem, neste caso, parte integrante do respetivo ato" (itálico nosso).

Num plano concreto do procedimento tributário, consagra o artigo 77.º, n.º 1, da Lei Geral Tributária, que "[a] decisão de procedimento é sempre fundamentada por meio de sucinta exposição das razões de facto e de direito que a motivaram"[27].

Pelo que uma decisão tomada num procedimento tributário cujo conteúdo seja exarado com recurso a sistemas de inteligência artificial não pode deixar de ser um ato devidamente fundamentado, expondo as razões de facto e de direito que motivaram a decisão final, não podendo o contribuinte ser o destinatário de um ato cujos contornos factuais e legais não conhece ou não lhe é permitido entender.

Não restará dúvidas que o tribunal competente será chamado a controlar se o ato tributário emanado permite conhecer o *iter cognitivo* para o desfecho do mesmo, ainda que este tenha sido emanado com suporte a sistemas de inteligência artificial, sendo reforçada esta última ideia na explicabilidade do próprio algoritmo que levou à tomada da decisão.

III. *Enviesamentos e discriminações*

Enraizado na premissa *ius aequalitatis*, o princípio da igualdade é um valor matricial consagrado no artigo 13.º da Constituição da República Portuguesa, sendo desenvolvido concretamente no domínio do procedimento tributário no artigo 55.º da Lei Geral Tributária. Consagra este último preceito legal que "a administração tributária exerce as suas atribuições na prossecução do interesse público, de acordo com os princípios da legalidade, *da igualdade*, da proporcionalidade, da justiça, *da imparcialidade* e da celeridade, no respeito pelas garantias dos contribuintes e demais obrigados tributários" (itálicos nossos). No contexto do procedimento tributário, o princípio da igualdade traduz-se no dever da Administração Tributária em tratar todos os obrigados tributários de forma equitativa, sem discriminação, seja na interpretação e aplicação das normas tributárias, seja na execução dos atos administrativos relacionados com a tributação. Isso implica que situações idênticas devem ser tratadas de forma idêntica e que qualquer tratamento diferenciado entre contribuintes deve estar objetivamente fundamentado.

Ora, como se tem vindo a sublinhar ao longo de todo o presente trabalho, a mera integração de sistemas de inteligência artificial no domínio do procedimento tributário não está dispensada das garantias estabelecidas e que se tradu-

[27] Joaquim Freitas da Rocha, *Lições de Procedimento e Processo Tributário*, 8.ª ed., Almedina, Coimbra, 2021, p. 126.

zem em verdadeiros direitos dos obrigados tributários, principalmente quando estamos perante atuações legitimadas segundo um parâmetro de discricionariedade variável. Por isso, os sistemas de inteligência artificial devem ser desenvolvidos e treinados de forma a não resultarem discriminações injustas e sem qualquer fundamentação que legitime a tomada de posição discriminatória[28].

Como tivemos ocasião de referir noutro momento, "[é] que um dos riscos das tecnologias de inteligência artificial é, precisamente, a discriminação algorítmica, em consequência dos enviesamentos introduzidos pela conceção do algoritmo ou dos resultados produzidos pelo seu funcionamento, que no domínio do estudo constante e de replicação são capazes de resultar em categorizações dos contribuintes ou em decisões sobre os contribuintes com critérios injustificáveis, discriminatórios e arbitrários"[29]. Falamos de uma discriminação que poderá resultar em atuações potencialmente lesivas para os contribuintes, tal como a produção de decisões atentatórias dos valores principiológicos da igualdade e não discriminação (*v.g.*, revogação ou atribuição de benefícios fiscais), principalmente quando a Administração atua com certo grau de discricionariedade[30].

A mera categorização negativa de um contribuinte em detrimento de outro poderá gerar atuações enviesadas e que são verdadeiramente atentatórias da igualdade tributária. Por isso, a adoção de sistemas inteligentes deve assegurar aos contribuintes que as normas jurídicas são aplicadas uniformemente em todas as decisões; que a política de decisão foi completamente definida (e essa escolha de política foi registada de forma segura) antes de serem conhecidos os casos específicos a serem decididos, reduzindo a possibilidade de manipular o procedimento para prejudicar uma pessoa em particular; que cada decisão possa ser reproduzida com base na política de decisão especificada e nos dados utilizados para essa decisão; e que, se uma decisão exigir quaisquer entradas de fatores

---

[28] Felipe Alonso Murillo, «Los Derechos Fundamentales como Límites al Empleo de Inteligencia Artificial en los Procedimientos Tributarios», in *Inteligencia Artificial y Administración Tributaria: Eficiencia Administrativa y Defensa de los Derechos de los Contribuyentes*, Thomson Reuteurs, Navarra, 2021, p. 65.

[29] Luís Manuel Pica, *A Inteligência Artificial no Direito Tributários*…, *op. cit.*, p. 89.

[30] Veja-se que o Grupo de Peritos da Comissão Europeia sobre Inteligência Artificial afirma que, "num contexto de IA, a igualdade implica que as operações do sistema não podem gerar resultados injustamente tendenciosos (p. ex., os dados utilizados para treinar os sistemas de IA devem ser o mais inclusivos possível, representando diferentes grupos da população). Tal exige que as pessoas e os grupos potencialmente vulneráveis, tais como trabalhadores, mulheres, pessoas com deficiência, minorias étnicas, crianças, consumidores, ou outras pessoas em risco de exclusão, sejam devidamente respeitados". Grupo Independente de Peritos de Alto Nível sobre a Inteligência Artificial da Comissão Europeia, *Orientações Éticas para uma IA de Confiança*, Bruxelas, 2019, p. 13.

que determinem a escolha e seleção aleatória de contribuintes, essas entradas estejam fora do controlo de qualquer das partes envolvidas[31]/[32].

5. Conclusões

A discussão aprofundada do princípio da legalidade, da tipicidade fiscal e do poder discricionário da Administração Tributária em Portugal, especialmente à luz dos desafios colocados pela inteligência artificial, revela a complexidade e a delicadeza de um equilíbrio fundamental no Estado de Direito.

Reafirmámos que o princípio da legalidade, ancorado no artigo 103.º, n.º 2, da Constituição da República Portuguesa e no artigo 8.º da Lei Geral Tributária, constitui uma reserva de lei, exigindo uma densidade normativa para os elementos essenciais do imposto. Esta reserva é um garante da previsibilidade e segurança jurídica para os contribuintes, protegendo-os contra a arbitrariedade do poder executivo.

Por outro lado, o poder discricionário da Administração Tributária, que surge precisamente da necessidade de interpretar e aplicar esses conceitos abertos segundo uma certa liberdade funcional normativamente delimitada, não é um poder arbitrário. A atuação da Administração Tributária está estritamente vinculada aos limites impostos pela lei e pelos princípios constitucionais da igualdade, proporcionalidade, justiça, imparcialidade e boa-fé, consagrados no artigo 55.º da Lei Geral Tributária e no artigo 266.º da Constituição da República Portuguesa. A fiscalização judicial dos atos administrativos tributários é a

---

[31] Oshua A. Kroll Joanna Huey; Solon Barocas; Edward W. Felten; Joel R. Reidenberg; David G. Robinson e Harlan Yu, «Accountable algorithms», *University of Pennsylvania Law Review*, Vol. 165, 2017, p. 657, disponível em https://scholarship.law.upenn.edu/penn_law_review/vol165/iss3/3/ [consultado em 09.08.2025].

[32] É por isso que o atual Regulamento Inteligência Artificial, aprovado pelo Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, exige que as empresas realizem auditorias e testes de impacto para identificar e mitigar possíveis vieses e discriminações criadas pelos sistemas de inteligência artificial. Para além disto, o regulamento também proíbe o uso de sistemas de inteligência artificial considerados inaceitáveis por serem, desde a sua conceção intrínseca, arbitrários ou discriminatórios, designadamente quando incluem sistemas de "pontuação social" que possam discriminar ou excluir indivíduos com base em avaliações automáticas de comportamento ou personalidade. Contudo, torna-se importante mencionar que o Regulamento não contém, objetivamente, uma disposição normativa sobre o "Proibição da Arbitrariedade Algorítmica". No entanto, o regulamento prevê garantias sólidas que se desenvolvem através deste princípio matricial. Em várias disposições legais do Regulamento é possível estabelecer artigos que tratam da transparência, não discriminação e requisitos para os sistemas de inteligência artificial de alto risco (*v.g.*, artigos 5.º, 13.º, 14.º, 15.º e 54.º). Esses artigos formam a base de uma estrutura regulatória proposta para garantir que os sistemas de inteligência artificial sejam utilizados de maneira ética, justa e não discriminatória na União Europeia, procurando estabelecer garantias sólidas contra a arbitrariedade algorítmica na utilização destes sistemas informáticos.

salvaguarda última contra o desvio de poder e a violação dos direitos dos contribuintes.

Finalmente, a irrupção da inteligência artificial na Administração Tributária impõe uma reflexão crítica sobre os limites funcionais da automação. Embora a inteligência artificial ofereça um potencial transformador em termos de eficiência e combate à fraude, a sua aplicação na tributação não pode prescindir da ponderação humana. A interpretação de conceitos tendencialmente abertos, de atuações com liberdade discricionária, o exercício de juízos de valor e a necessidade de garantir a justiça material no caso concreto são tarefas que, por inerência, exigem a sensibilidade, o contexto e a capacidade de análise ética que apenas o ser humano possui. A inteligência artificial deve ser um instrumento de apoio à decisão, um facilitador para a Administração Tributária, mas nunca o decisor final em matérias que afetam tão profundamente os direitos e deveres dos cidadãos. A transparência dos algoritmos, a possibilidade de controlo humano efetivo (*human-in-the-loop*) e a clara responsabilização pelas decisões são pilares indispensáveis para que a integração da inteligência artificial no sistema fiscal se faça no respeito pelos princípios do Estado de Direito.

Em suma, o desafio reside em conciliar a eficiência e a modernização tecnológica com a intangível, mas fundamental dimensão garantística do Direito Fiscal. A legalidade e a tipicidade, mesmo que abertas, e a discricionariedade, mesmo que limitada, são essenciais para um sistema fiscal justo e democrático, onde a inteligência da máquina serve a justiça administrada pelo Homem.

DA ILUSÃO DIGITAL DAS DEEPFAKES À REALIDADE PENAL: A TUTELA DA AUTODETERMINAÇÃO SEXUAL EM MACAU

*Ng Chan Fai**

Resumo: O estudo analisa, sob uma perspectiva dogmático-penal e comparada, o impacto das tecnologias *deepfake* na tutela da autodeterminação sexual e na definição jurídica do consentimento, tomando como referência o ordenamento penal da Região Administrativa Especial de Macau (RAEM). O Código Penal de Macau consagra, no capítulo relativo aos crimes contra a liberdade e a autodeterminação sexual (artigos 157.º a 173.º), as formas clássicas de ofensa à liberdade sexual, como a violação, a coacção e o abuso sexual. Contudo, as agressões virtuais de natureza sexual escapam a esta tipificação e revelam lacunas na protecção penal da dignidade e da liberdade sexual.

* Doutor em Direito pela Universidade do Minho, Mestre em Direito pela Universidade de Macau e Licenciado em Direito pela Universidade Católica Portuguesa. É Assessor Jurídico na Direcção dos Serviços de Administração e Função Pública da RAEM e docente a tempo parcial na Faculdade de Direito da Universidade de Macau e na Universidade Politécnica de Macau. Trabalhou no Gabinete do Presidente do Tribunal de Última Instância, onde coordenou a redacção bilingue e a compilação oficial de jurisprudência. É Investigador Integrado do JusGov – Centro de Investigação em Justiça e Governação (Braga, Portugal). E-mail: ngchanfai@hotmail.com.

A investigação identifica desafios que a manipulação digital coloca à prova penal – desde a recolha e autenticação da prova electrónica até à valoração pericial – e evidencia a insuficiência do regime vigente para garantir tutela efectiva às vítimas. Com base na análise comparada das soluções adoptadas em Portugal, na União Europeia, na China, em Hong Kong e nos Estados Unidos, defende-se a criação de um tipo penal autónomo para a pornografia *deepfake*, neutro quanto à tecnologia e centrado na ausência de consentimento e na identificabilidade da vítima. Conclui-se que a autodeterminação sexual constitui o bem jurídico central e directamente lesado, enquanto a honra e a privacidade da vítima são bens jurídicos conexos, tutelados de forma reflexa no quadro da dignidade sexual. Esta delimitação reforça a coerência sistemática da análise dogmático-penal e da política criminal proposta, exigindo uma política criminal prospectiva que concilie inovação tecnológica e protecção dos direitos fundamentais.
Palavras-chave: *Deepfake*; Autodeterminação sexual; Consentimento; Dignidade sexual; Tutela penal; Crimes contra a liberdade sexual; Pornografia não consentida; Prova digital; Política criminal; Direito Penal de Macau.

Abstract: This study analyses, from a dogmatic and comparative criminal law perspective, the impact of deepfake technologies on the protection of sexual self-determination and on the legal concept of consent, taking Macao's criminal law as its reference. The Macao Penal Code, in the chapter on crimes against liberty and sexual self-determination (Articles 157 to 173), codifies the traditional offences of rape, sexual coercion, and sexual abuse. However, virtual sexual assaults fall outside this typification, exposing structural gaps in the criminal protection of dignity and sexual freedom.

The research identifies the challenges that digital manipulation poses to criminal evidence – from the collection and authentication of electronic data to its forensic evaluation – and highlights the inadequacy of the current framework in ensuring effective victim protection. Drawing on comparative analysis of legal approaches in Portugal, the European Union, Mainland China, Hong Kong, and the United States, it argues for the creation of an autonomous criminal offence for deepfake pornography, technologically neutral and centred on the absence of consent and the identifiability of the victim. It concludes that sexual self-determination should be recognised as the legal interest directly violated, requiring a forward-looking criminal policy capable of reconciling technological innovation with the protection of fundamental rights.

Keywords: Deepfake; Sexual self-determination; Consent; Sexual dignity; Criminal protection; Crimes against sexual liberty; Non-consensual pornography; Digital evidence; Criminal policy; Macao Criminal Law.

1. Introdução

Nos últimos anos, o desenvolvimento das tecnologias *deepfake* baseadas em inteligência artificial tem evoluído a um ritmo particularmente acelerado. Por intermédio de técnicas de *aprendizagem profunda*, tornou-se possível gerar imagens visuais e sons dotados de um realismo notável. Embora tais tecnologias apresentem utilizações legítimas e socialmente benéficas – designadamente na produção cinematográfica ou na criação de conteúdos pedagógicos –, a sua apropriação indevida tem vindo a produzir efeitos negativos de crescente gravidade. Estudos recentes demonstram que mais de 90% dos vídeos *deepfake* actualmente disponíveis na Internet correspondem a representações pornográficas criadas sem o consentimento das pessoas visadas, sendo cerca de 90% dessas dirigidas contra mulheres[1].

Recorrendo a esta tecnologia, agentes mal-intencionados conseguem, sem qualquer autorização, fabricar vídeos eróticos simulados, de carácter hiper-realista, nos quais o rosto da vítima é digitalmente inserido em material pornográfico preexistente. Tais conteúdos pornográficos *deepfake* não consentidos constituem uma grave violação da integridade psicológica, da reputação e, sobretudo, do direito fundamental à autodeterminação sexual das vítimas.

O presente estudo propõe-se examinar, sob uma perspectiva dogmático-penal e comparada, os desafios que este fenómeno coloca ao conceito jurídico de "consentimento sexual" e aos mecanismos de tutela da autodeterminação sexual, analisando as insuficiências do regime penal vigente em Macau e delineando, por conseguinte, possíveis respostas de natureza substantiva e processual no âmbito do Direito Penal.

Embora o estudo se centre no ordenamento jurídico de Macau, as questões analisadas – em especial as relativas à inteligência artificial e às tecnologias *deepfake* – inserem-se nos desafios transversais que o avanço tecnológico vem colocando ao Direito Penal contemporâneo, coincidindo, em termos substantivos, com as linhas de investigação e os eixos temáticos do grupo E-Tec, dedicados à inteligência artificial, à robótica e às tecnologias disruptivas para o Direito.

---

[1] Centro Europeu de Direito Público, *Relatório sobre a Violência Digital 2023*, Centro Europeu de Direito Público, 2023 – O relatório aponta que, entre 2019 e 2023, o número de casos envolvendo imagens *deepfake* aumentou em 550%, sendo que 99% das vítimas eram mulheres.

2. O impacto das tecnologias *deepfake* sobre o consentimento sexual e a autodeterminação sexual

O aparecimento de conteúdos pornográficos gerados por *deepfake* veio introduzir um novo desafio ao conceito tradicional de "consentimento" no Direito Penal. A autodeterminação sexual, enquanto direito de cada indivíduo de decidir livremente se, quando e de que modo deseja praticar actos de natureza sexual, constitui o núcleo essencial da valoração jurídico-penal dos crimes de violação e de coacção sexual. Todavia, nas imagens artificialmente geradas, a vítima jamais participou na acção retratada, sendo, no entanto, forçada a "comparecer" digitalmente em cenas de cariz sexual. Nestas circunstâncias, inexiste qualquer consentimento efectivo, e nem sequer a possibilidade de recusa, verificando-se, por conseguinte, uma violação digital do seu direito à autodeterminação sexual.

O carácter verosímil das imagens leva o público a acreditar que a pessoa representada agiu voluntariamente, originando o fenómeno do chamado "consentimento fictício" ou "consentimento aparente" – uma simulação de vontade que, na realidade, nunca existiu. Este artifício causa às vítimas profundos danos psicológicos e reputacionais, ao mesmo tempo que perturba a clareza dos critérios jurídicos de distinção entre actos praticados "com consentimento" e aqueles cometidos "contra a vontade".

As *deepfakes* pornográficas não afectam apenas os direitos clássicos à honra e à privacidade, mas implicam igualmente uma usurpação da identidade e da dignidade sexual da pessoa. Vários autores descrevem esta situação como um verdadeiro "sequestro da identidade sexual" da vítima, inserindo-a no quadro mais amplo do abuso sexual imagético e da violência sexual facilitada por tecnologia[2], já que o agressor manipula, sem autorização, a sua imagem e o seu corpo digitalizado. Ainda que o corpo exposto não corresponda ao real, o domínio da intimidade sexual é violado. Do ponto de vista da tutela penal dos bens jurídicos, as *deepfakes* sexuais atentam contra a autodeterminação e a intimidade sexual – bens jurídicos centrais –, configurando uma nova forma de lesão da liberdade e dignidade sexual, em linha com a qual tem vindo a ser conceptualizada a violação da privacidade sexual como atentado directo à autonomia sexual da pessoa[3].

---

[2] Clare McGlynn, «The 'New Voyeurism': Criminalising the Creation of Pornographic Deepfakes», *Journal of Law and Society*, 2025 (online first); Rüya Tuna Toparlak, «Criminalising Deep Fake Pornography: A Gender-Specific Analysis of Image-Based Sexual Abuse», *cognitio 2023/1*, Betriebs-Berater (Strafrecht), 2023.

[3] V. Danielle Keats Citron, «Sexual Privacy», *Yale Law Journal*, 128(7), 2019, pp. 1870–1961; *idem*, *William & Mary Law Review*, Vol. 62, Issue 6, 2021, pp. 1765–1834; Alexander De Ruiter, «The Distinct Wrong of Deepfakes», *Ethics and Information Technology*, 25, 2023, pp. 1–12.

De acordo com estudos de auditoria realizados em 2019 sobre plataformas públicas[4], cerca de 96% dos vídeos *deepfake* identificados correspondiam a conteúdos pornográficos, praticamente todos tendo mulheres como alvo principal. Em 2024, uma revisão académica reafirmou que, não obstante certas limitações metodológicas das fontes disponíveis, o "*deepfake* pornográfico" continua a representar a modalidade de abuso mais recorrente e sistemática[5].

3. Insuficiências do regime penal vigente em Macau

Os artigos 157.º a 173.º do Código Penal de Macau já tipificam os crimes sexuais tradicionais, como a violação, a coacção sexual, o acto sexual mediante fraude, o abuso sexual, os actos de natureza libidinosa e a produção ou comercialização de material pornográfico[6]. Todavia, as imagens sexuais virtuais não consentidas, criadas através de tecnologias *deepfake*, não se enquadram expressamente em nenhum destes tipos legais.

Com efeito, tal lacuna não é exclusiva de Macau: diversos ordenamentos jurídicos ainda não lograram abranger de modo tempestivo esta nova forma de ofensa, deixando as vítimas, mesmo quando gravemente lesadas, numa situação de verdadeira desproteção jurídica. No caso específico de Macau, o conteúdo pornográfico *deepfake* não envolve a participação real da vítima em qualquer acto sexual, o que torna inviável a aplicação das figuras de violação ou de abuso sexual. O agente não recorre à coacção nem à fraude para obter um contacto físico, e, portanto, o "acto sexual contra a vontade" – núcleo dos crimes sexuais clássicos – não se consuma[7].

Contudo, a humilhação, a degradação da personalidade e a violação da autodeterminação sexual que resultam da difusão dessas imagens são, na sua natureza, substancialmente análogas à violência ou ao abuso sexual. Esta situação revela um vazio normativo: a dificuldade de enquadrar juridicamente, sob os

---

[4] Henry Ajder; Giorgio Patrini; Francesco Cavalli e Laurence Cullen, *The State of Deepfakes: Landscape, Threats, and Impact*, Deeptrace, setembro de 2019 – O relatório indica que cerca de 96% dos vídeos *deepfake* têm conteúdo pornográfico. Pode ver-se o relatório em formato PDF reproduzido em: https://docslib.org/doc/12559428/the-state-of-deepfakes-landscape-threats-and-impact.

[5] Alena Birrer; Natascha Just, «What we know and don't know about deepfakes: An investigation into the state of the research and regulatory landscape», *New Media & Society*, 2024, disponível em: https://journals.sagepub.com/doi/full/10.1177/14614448241253138.

[6] Código Penal da Região Administrativa Especial de Macau, aprovado pela Lei n.º 9/1995, com as alterações introduzidas pela Lei n.º 8/2017 (Boletim Oficial da RAEM, I Série, n.º 27, 3 Julho 2017), artigos 1.º, 157.º-173.º.

[7] Código Penal Português, aprovado pelo Decreto-Lei n.º 400/82, de 23 de Setembro, com as alterações subsequentes, artigos 163.º (Coacção sexual) e 164.º (Violação).

requisitos típicos do crime e sob o regime de imputação penal, as condutas relacionadas com simulações digitais e com o chamado "consentimento fictício".

Sob a perspectiva da tutela dos bens jurídicos, as *deepfakes* pornográficas atentam simultaneamente contra a autodeterminação sexual, a privacidade e a honra[8]. Embora estes bens sejam indiscutivelmente essenciais, o ordenamento vigente carece de tipos legais específicos que permitam sancionar de modo eficaz tais comportamentos, gerando uma clara disfunção na protecção dos direitos das vítimas.

O regime penal de Macau continua a centrar-se predominantemente em condutas ocorridas em contextos reais[9], como a produção ou a difusão de conteúdos de natureza sexual envolvendo menores, punível com pena até cinco anos de prisão[10]. No entanto, quando o material em causa diz respeito a adultos e é integralmente gerado por inteligência artificial, a situação já não se insere na protecção especial da infância, nem se adequa ao tipo legal de "produção de material pornográfico".

Por outro lado, os crimes de difamação ou injúria poderiam, em tese, ser invocados, mas estes tutelam primordialmente a veracidade de factos ou declarações ofensivas, não abrangendo necessariamente a dimensão de humilhação sexual associada a imagens artificiais, como tem sido salientado na doutrina que analisa a inadequação dos remédios clássicos de difamação face à pornografia *deepfake*[11].

Conclui-se, portanto, que o ordenamento penal macaense evidencia uma lacuna estrutural de tipicidade, na medida em que o uso não consentido da ima-

---

[8] Diretiva (UE) 2024/1385 do Parlamento Europeu e do Conselho, de 12 Junho 2024, relativa à criminalização das *deepfakes* de natureza sexual e à cooperação judicial no domínio do crime cibernético [2024] JOUE L 202/1.

[9] Túlio Felippe Xavier Januário, «Artificial Intelligence and Image-Based Sexual Abuse: An Analysis of Non-Consensual Deepfakes in Light of Portuguese Criminal Law», *Revista Internacional de Direito Penal*, 95(2), 2024, p. 483.

[10] Região Administrativa Especial de Macau, *Código Penal de Macau*, promulgado pelo Decreto-Lei n.º 58/95/M, de 14 de Novembro de 1995, em vigor desde 1 de Janeiro de 1996 (Boletim Oficial n.º 46/1995, p. 2300), disponível em bo.io.gov.mo. Posteriormente alterado pela Lei n.º 8/2017 (Boletim Oficial n.º 26, pp. 614-619), que aditou o artigo 170.º-A (Pornografia de menor), o qual proíbe expressamente a produção, comercialização, importação, exportação ou difusão de imagens pornográficas envolvendo menores, punindo tais condutas com pena até oito anos de prisão, e prevendo pena até três anos mesmo para a mera retransmissão, refletindo assim uma política de "tolerância zero". Ver igualmente Gabinete de Comunicação Social do Governo da RAEM, *Revisão do Código Penal – Breve Apresentação*, Comunicado de Imprensa, 2017.

[11] V. Abigail George, «Defamation in the Time of Deepfakes», *Columbia Journal of Gender and Law*, Vol. 43, No. 2, 2024, pp. 1–52; Serviço de Estudos do Parlamento Europeu, «Tackling Deepfakes in European Policy», European Parliamentary Research Service (EPRS), 2021, pp. 13–16.

gem de outrem para fins libidinosos ou pornográficos não constitui, à presente data, objecto de proibição penal expressa. Nem o capítulo dos crimes contra a autodeterminação sexual, nem o dos crimes contra a honra se mostram aptos a abranger adequadamente esta nova forma de agressão à liberdade e à dignidade sexual.

3.1. *Lacunas dogmático-penais na tutela da autodeterminação sexual virtual*

A experiência recente demonstra que as *deepfakes* de cariz sexual lesam, antes de mais, a autodeterminação sexual enquanto bem jurídico central dos crimes sexuais, porquanto suprimem a faculdade de decidir *se, quando e como* alguém é representado em contexto sexual[12]. Tal lesão ocorre sem acto físico nem contacto, mas mediante simulação digital hiper-realista que projecta no espaço público um *consentimento que nunca existiu* (*consentimento fictício*). Deste modo, a agressão incide directamente sobre a liberdade sexual negativa, além de repercutir-se na honra e na privacidade.

Propõe-se, assim, a categoria teórica de "acto de agressão sexual virtual", funcionalmente equivalente – no plano da tutela penal – às formas clássicas de agressão, apesar da ausência de corporeidade[13].

Nesta senda, importa indagar se e em que termos tal agressão virtual poderia ser absorvida pelos tipos penais existentes em Macau (artigos 157.º-173.º do Código Penal), ou se, ao invés, revela uma lacuna típica que exige novo tipo legal. Como assinala Maria João Antunes Pereira, no contexto do direito português, existe actualmente uma *lacuna legal* em matéria de *deepfakes* sexuais envolvendo adultos, não coberta pelos crimes sexuais tradicionais[14].

(i) Violação e coacção sexual (artigos 163.º e 164.º do Código Penal português):

Em primeiro lugar, os tipos nucleares de violação e de coacção sexual pressupõem, dogmaticamente, a realização de actos sexuais corpóreos obtidos por violência, ameaça grave ou outros meios de constrangimento sobre a víti-

---

[12] Clare McGlynn; Rüya T. Toparlak, «Criminalising the Creation of "Deepfake Porn": A New Frontier of Image-Based Sexual Abuse», SSRN Working Paper SSRN, 2024, disponível em: https://papers.ssrn.com/sol3/Delivery.cfm/4894256.pdf [acedido em 08/10/2025].

[13] *Ibidem*.

[14] Maria João Antunes Pereira, *Comentário ao Código Penal: Parte Especial – Crimes Contra a Liberdade e Autodeterminação Sexual*, Almedina, Coimbra, 2018.

ma[15]. O modelo português (artigos 163.º e 164.º do Código Penal) e o correspondente de Macau (artigos 157.º e ss. do Código Penal de Macau, redacção da Lei n.º 8/2017) partilham esta estrutura assente em condutas que culminam em actos sexuais reais. Ora, nas *deepfakes*, não há acto sexual praticado pela vítima, mas uma representação simulada da sua sexualidade. Falta, pois, o elemento nuclear do tipo – a realização de um acto sexual *contra a vontade* mediante constrangimento físico ou equivalente. Qualquer tentativa de enquadrar a simulação digital nessa noção redundaria numa analogia *in malam partem*, vedada pelo princípio da legalidade penal[16].

(ii) "Coacção indirecta" pela simulação digital:

Por outro lado, poder-se-ia sustentar que a difusão de uma imagem *deepfake* sexual cria uma pressão social (exposição pública, ostracismo, risco profissional) capaz de restringir a autodeterminação futura da vítima – uma forma de coacção indirecta exercida sem contacto físico, mas por intermédio de *constrangimento simbólico*. Ainda que tal leitura capte a equivalência material do desvalor (o sequestro simbólico da identidade sexual), o nexo típico dos crimes de coacção/violação continua a exigir uma relação de domínio efectivo do agente sobre a vítima tendente à obtenção de um acto sexual real. Assim, esta recondução deve ser vista apenas em tese (hipótese interpretativa), pois a subsunção literal seria juridicamente insustentável.

(iii) Finalmente, quanto ao acto sexual mediante fraude e a (in)viabilidade de *analogia in bonam partem*, o ordenamento da RAEM prevê o acto sexual mediante fraude entre os tipos que exigem a participação da vítima em acto sexual real viciado por erro. Nas *deepfakes*, ocorre o inverso: não há acto sexual da vítima, apenas o fabrico de uma aparência de consentimento (*consentimento fictício*). Tentativas de incluir tal simulação digital neste tipo, invocando *analogia in bonam partem*, não resolvem o problema, pois (a) esta analogia só opera a favor do arguido e (b) aqui implicaria uma extensão em sentido incriminador, proibida pelo princípio da legalidade[17]. Esta conclusão é de ordem dogmática e não assenta em jurisprudência consolidada.

---

[15] Jorge de Figueiredo Dias, *Direito Penal: Parte Geral*, vol. I, Coimbra Editora, Coimbra, 2007.

[16] Lisbon Public Law Centre, «A Criminalização de Deepfakes Sexuais: Um Novo Desafio para o Direito Penal Digital», in *Lisbon Public Law Blog*, 2024, disponível em: https://lisbonpubliclaw.pt/en/blog/a-criminalizacao-de-deepfakes-sexuais-um-novo-desafio-para-o-direito-penal-digital/ [acedido em 08/10/2025].

[17] Reuters, «UK to Make Sexually Explicit Deepfakes a Criminal Offence», in *Reuters Legal Section*, 7 January 2025, disponível em: https://www.reuters.com/world/uk/britain-make-sexually-explicit-deepfakes-crime-2025-01-07/ [acedido em 08/10/2025].

(iv) Consequência sistemática: lacuna típica e necessidade de tipificação autónoma: Do ponto de vista dogmático-penal, as *deepfakes* sexuais lesam o mesmo bem jurídico protegido pelos crimes sexuais clássicos – a autodeterminação sexual –, mas por um meio desmaterializado que não encontra correspondência nos elementos objectivos dos tipos existentes. A tentativa de recondução via "coacção indirecta" ou "fraude" gera uma tensão sistemática com os requisitos de corporeidade e com a proibição de analogia incriminadora. Conclui-se, pois, em termos teóricos, que se trata de uma lacuna típica estrutural: o ordenamento tutela o material pornográfico envolvendo menores, bem como a autodeterminação sexual – bem jurídico central dos crimes sexuais – e, de modo reflexo, a honra e a privacidade, mas não contém tipo penal que proíba expressamente a produção e difusão de representações sexuais não consentidas geradas por inteligência artificial. Tal constatação aponta para a necessidade de uma tipificação autónoma, tecnologicamente neutra e centrada na ausência de consentimento e na identificabilidade da vítima – em sintonia com as tendências emergentes na União Europeia, como ilustrado pela Diretiva (UE) 2024/1385, que estabelece padrões mínimos de criminalização para a criação e difusão de representações sexuais não consentidas geradas por inteligência artificial, sem prejuízo da autonomia normativa do ordenamento jurídico de Macau[18].

Nota metodológica: as categorias de "acto de agressão sexual virtual", "coacção indirecta" e a tese da "lacuna típica estrutural" são propostas teóricas apresentadas por V. Danielle Citron. Embora encontrem eco em análises comparativas[19], não existe, até à data, jurisprudência confirmada em Macau ou Portugal que reconheça expressamente estas condutas como crimes sexuais tradicionais.

---

[18] Diretiva (UE) 2024/1385 do Parlamento Europeu e do Conselho, de 12 Junho 2024, [2024] JOUE L 202/1; Boletim Oficial da RAEM, Lei n.º 8/2017 – Alteração ao Código Penal (I Série, n.º 27, 3 Julho 2017). Comparativamente, a Diretiva serve apenas como referência de política criminal, sem aplicação directa em Macau.

[19] Rüya Tuna Toparlak, «Criminalising Deep Fake Pornography…», *op. cit.*; Clare McGlynn, «The 'New Voyeurism': Criminalising the Creation of Pornographic Deepfakes», *op. cit.*

4. Desafios na obtenção e valoração da prova digital em casos de *deepfake*

Para além das insuficiências de natureza substantiva, os casos relacionados com *deepfakes* enfrentam igualmente sérios desafios no domínio processual, designadamente no que respeita à recolha, conservação e perícia da prova digital.

O Código de Processo Penal de Macau contém disposições específicas sobre a intercepção de comunicações, a apreensão de dados electrónicos e a conservação de informação digital, sendo complementado pela Lei n.º 11/2009, que institui um regime próprio de obtenção de prova electrónica. A reforma legislativa de 2020 veio ampliar os poderes de investigação em matéria de criminalidade informática[20], permitindo, entre outras medidas, o acesso a dados armazenados em servidores localizados fora do território da RAEM. Assim, mediante autorização judicial, as autoridades podem ordenar ao suspeito a entrega das credenciais de acesso às suas contas de serviços em nuvem (*cloud services*), facultando o acesso directo, a partir dos seus dispositivos, às informações alojadas em servidores estrangeiros[21].

Esta inovação criou, em termos teóricos, um canal mais ágil para a obtenção de prova em investigações de natureza digital, reduzindo a dependência dos mecanismos clássicos de cooperação judiciária internacional. Todavia, no que toca especificamente ao tratamento probatório de conteúdos *deepfake*, tanto a legislação vigente como a experiência prática revelam-se ainda insuficientes.

A identificação e a autenticação de vídeos gerados por inteligência artificial apresentam um elevado grau de complexidade técnica[22]. Actualmente, não existe em Macau qualquer norma que defina critérios objectivos para determinar se uma gravação foi artificialmente manipulada. Os peritos forenses digitais devem recorrer a algoritmos e ferramentas de análise avançada – nomeadamente à detecção de anomalias nos *pixels*, à verificação de padrões de metadados e à

---

[20] Lei n.º 4/2020 (que altera a Lei n.º 11/2009), publicada no Boletim Oficial da Região Administrativa Especial de Macau, disponível em: bo.io.gov.mo.

[21] Governo da Região Administrativa Especial de Macau, *Comunicado de Imprensa de 2019* – Expansão do âmbito da recolha de provas em linha e explicação das políticas relativas à autorização judicial. Em 2019, o Governo propôs a revisão da lei sobre cibercriminalidade, prevendo autorizar a polícia, mediante autorização judicial, a aceder a dados armazenados em servidores na nuvem situados no estrangeiro, bem como criar os crimes de exploração de falsas estações base e de desobediência qualificada, com o objectivo de reforçar o combate à criminalidade informática transfronteiriça. Disponível em: https://www.gov.mo/pt/noticias/215251/.

[22] Shavez Mushtaq Qureshi et al., «Deepfake forensics: a survey of digital forensic methods for multimodal deepfake identification on social media», *PeerJ Computer Science*, 10:e2037, 2024; Irene Amerini et al., «Deepfake Media Forensics: Status and Future Challenges», *Journal of Imaging*, 11(3), 2025, p. 73.

identificação de marcas digitais ou *watermarks* –, a fim de avaliar a autenticidade da prova, em consonância com o que vem sendo identificado na literatura forense internacional como um dos principais desafios probatórios na era dos *deepfakes*.

Todavia, a limitação técnico-pericial das autoridades macaenses não só evidencia a necessidade de uma reforma estrutural da perícia digital, como também fragiliza a capacidade do sistema judicial para enfrentar a crescente sofisticação das técnicas de falsificação digital.

As autoridades judiciais reconhecem que os infractores recorrem cada vez mais ao armazenamento em nuvem para ocultar provas, o que eleva substancialmente a dificuldade da investigação e da recolha de elementos probatórios. No primeiro caso registado em Macau envolvendo fraude com recurso a *deepfakes*, um grupo criminoso fabricou vídeos falsificados através de substituição facial e imitação vocal, associando-os a páginas *phishing* destinadas a enganar as vítimas. A dispersão das provas por várias plataformas comprometeu a eficácia dos métodos tradicionais de investigação[23].

A Polícia identificou também um vídeo manipulado que usurpava a imagem do Chefe do Executivo, Sam Hou Fai, aparentando promovê-lo como representante de uma plataforma de investimento fictícia. O conteúdo foi difundido nas redes sociais para atrair o público. Embora não tenham surgido vítimas formais, a Polícia emitiu um alerta, advertindo contra este tipo de burla baseada na apropriação indevida de imagem pública[24].

De igual modo, a MGM China declarou que terceiros criaram um vídeo *deepfake* apresentando falsamente a sua Presidente, Pansy Ho, a apoiar um alegado plano de investimento. A empresa afirmou tratar-se de um conteúdo

[23] «Macao police paying close attention to new trends in AI-related crime», Gabinete de Comunicação Social do Governo da RAEM, 28 de Maio de 2025 – A notícia refere que, em Abril de 2025, a polícia detectou o primeiro caso de burla envolvendo imagens *deepfake* geradas por inteligência artificial, utilizando técnicas de substituição facial e imitação de voz. Disponível em: https://www.gov.mo/en/news/359423/.

[24] «Police warn of deepfake video impersonating Sam Hou Fai in investment scam», *Macau Daily Times/Macau Post Daily*, 20 de Junho de 2025 – A notícia relata a circulação de um vídeo falsificado que utilizava a imagem de Sam Hou Fai para induzir investimentos fraudulentos, levando a polícia a emitir um alerta público. Disponível em: https://macaudailytimes.com.mo/police-warn-of-deepfake-video-impersonating-sam-hou-fai-in-investment-scam.html.

totalmente fraudulento, desvinculando-se dele e reservando-se o direito de agir judicialmente[25].

Estes casos revelam a vulnerabilidade estrutural do sistema probatório face à sofisticação tecnológica. A análise de conteúdos *deepfake* continua a depender de perícia técnica e de equipas especializadas, mas Macau ainda carece de recursos humanos e infra-estruturas adequadas. Sem esse reforço, será difícil acompanhar a crescente complexidade das fraudes digitais e garantir a fiabilidade da prova em tribunal.

A admissibilidade e a integridade da prova digital constituem desafios centrais nos processos envolvendo *deepfakes*. Quando as autoridades obtêm um vídeo suspeito difundido em linha, é indispensável comprovar a origem do ficheiro, assegurar a cadeia de custódia e demonstrar que a imagem representa efectivamente a vítima e foi criada sem o seu consentimento. Na ausência de orientações técnicas claras e de padrões periciais reconhecidos, a credibilidade desta prova pode ser seriamente questionada em tribunal.

A natureza volátil dos conteúdos *deepfake* agrava o problema: a sua replicação é quase instantânea e a eliminação do material original não impede a sua difusão contínua. Assim, as vítimas permanecem expostas enquanto os vestígios digitais se dispersam em múltiplas plataformas, dificultando o controlo probatório.

Embora as recentes reformas legais permitam, com autorização judicial, o acesso directo a dados armazenados em servidores estrangeiros através dos dispositivos do arguido, tal mecanismo depende da sua cooperação efectiva. Se o autor do crime estiver no exterior ou for desconhecido, as autoridades de Macau continuam dependentes de pedidos de cooperação internacional, notificações à INTERPOL ou acordos bilaterais de assistência judiciária.

Na prática, os mecanismos de colaboração transnacional em matéria de *deepfakes* e crimes sexuais digitais ainda são incipientes: os procedimentos são lentos e a recolha de prova carece de eficácia temporal. Consequentemente, mesmo com base legal adequada, as limitações técnicas e institucionais persistem, tornando a investigação e a acção penal nestes casos um verdadeiro campo de obstáculos.

---

[25] «Statement from MGM China», MGM China Holdings Limited – Comunicado oficial no qual a empresa declara que o vídeo *deepfake* em causa utilizou indevidamente a imagem de Ms. Pansy Ho, Presidente e Directora Executiva da MGM China Holdings Limited, e que tal conteúdo não tem qualquer relação com a empresa nem com a própria. A MGM China reafirma que reserva todos os direitos legais para agir contra a falsificação. Ver também o comunicado complementar «MGM China warns public of deepfake investment scam». Disponível em: https://en.mgmchinaholdings.com/media-releases?item=752&.

5. Outros países e regiões: experiências legislativas comparadas

Para responder aos desafios gerados pelas tecnologias *deepfake*, diversos ordenamentos jurídicos têm procurado, nos últimos anos, delinear estratégias legislativas adequadas. O presente estudo toma como referência os casos da China, de Hong Kong, de Portugal, dos Estados Unidos e da Alemanha, apresentando uma síntese das respectivas experiências normativas no tratamento das imagens sexuais virtuais e de conteúdos análogos a agressões sexuais simuladas, complementando-as com exemplos de outras jurisdições, com o objectivo de oferecer subsídios comparativos úteis à realidade de Macau.

5.1. *China Continental – Regime jurídico tripartido de governação das deepfakes*

- Na China continental, a utilização indevida de tecnologias de inteligência artificial para a criação de **deepfakes** representa uma forma de violação da autodeterminação sexual, na medida em que implica a apropriação, sem consentimento, de traços identificáveis de uma pessoa e a sua inserção em contextos sexualizados, privando-a do controlo sobre a sua própria imagem sexual. O actual enquadramento jurídico assenta em três eixos regulatórios – os direitos da personalidade, a protecção de dados pessoais e a supervisão cibernética – que, em conjunto, formam uma estrutura de governação, ainda que marcada por insuficiências quanto à tipicidade penal, à carga probatória e à cooperação transfronteiriça.

- No plano civil, o Livro das Personalidades do **Código Civil da República Popular da China** sistematiza os direitos à imagem, à reputação, à privacidade e à protecção da informação pessoal, fornecendo uma base directa de tutela às vítimas de pornografia **deepfake**. O artigo 1018.º reconhece o direito de cada pessoa a produzir, utilizar, divulgar ou autorizar o uso da sua imagem, definindo "imagem" como qualquer representação visual que permita identificar o indivíduo. Assim, a montagem digital de um rosto sobre um corpo em material sexual e a sua divulgação configuram o uso e a publicação não consentidos da imagem alheia, legitimando a vítima a exigir a cessação da violação, a eliminação

dos efeitos e a indemnização pelos danos sofridos[26]. O artigo 1010.º, por sua vez, introduz pela primeira vez a responsabilidade civil pelo assédio sexual, impondo às entidades públicas e privadas o dever de adoptar "medidas razoáveis" de prevenção e resposta, o que serve de base para enfrentar casos de assédio sexual **deepfake** em contextos laborais ou académicos[27]. O artigo 995.º estabelece ainda uma cláusula geral de tutela da personalidade, permitindo ao tribunal interpretar teleologicamente o bem jurídico da autodeterminação sexual, mesmo quando o conteúdo em causa não representa um acto "real", assegurando uma protecção material equivalente.

- No domínio da protecção de dados, a **Lei de Protecção de Informação Pessoal** classifica os dados biométricos como "informação pessoal sensível", exigindo um propósito específico, necessidade estrit[28] e consentimento separado para o seu tratamento. Este elevado limiar de protecção incide directamente sobre a matéria-prima das **deepfakes**, tornando ilícita, desde logo, a recolha, modelação e síntese não consentidas desses dados, e oferecendo às vítimas fundamentos para acção conjunta nas vias administrativa e civil.

- Na vertente de regulação cibernética, o **Regulamento sobre Serviços de Síntese Profunda na Internet**[29] impõe obrigações materiais aos fornecedores de serviços e às entidades de apoio técnico: qualquer edição, síntese ou geração de dados biométricos de pessoas naturais deve ser claramente identificada e precedida de consentimento expresso; conteúdos susceptíveis de confundir o público devem ser marcados de modo visível; e deve existir um sistema de "reclamação-tratamento-rastreabilidade-relato". As plataformas têm o dever de rever, remover e desmentir conteúdos ilegais, manter registos de logs e registar os algoritmos

---

[26] *Código Civil da República Popular da China*, aprovado em 2020 e em vigor desde 1 de Janeiro de 2021, artigo 1018.º.

[27] *Ibidem*, artigo 1010.º.

[28] *Lei da República Popular da China sobre a Proteção das Informações Pessoais*, aprovada em 20 de Agosto de 2021 e em vigor desde 1 de Novembro de 2021, artigo 28.º.

[29] Conhecida como *Regulamento sobre a Síntese Profunda* (　　), proíbe a criação de imagens de síntese profunda sem o consentimento da pessoa retratada e exige a indicação clara de que o conteúdo foi gerado por inteligência artificial, sendo considerada uma das normas mais rigorosas do mundo em matéria de controlo de *deepfakes*.

utilizados, reforçando simultaneamente a rastreabilidade e a responsabilização tanto na produção como na circulaçã[30]. O regulamento complementar sobre **Serviços de Inteligência Artificial Generativa** impõe, de modo articulado, obrigações quanto à segurança de conteúdos, identificação, tratamento de queixas e fiscalização posterior.

- Na prática judicial, os tribunais chineses já reconheceram que a apropriação de voz ou timbre identificável de actores de dobragem[31], ou o uso de software de troca facial, pode constituir violação da personalidade e dos dados pessoais. Estes precedentes consolidam três critérios centrais de prova: identificabilidade, imputabilidade e autenticidade verificável. Embora o sistema jurídico chinês não consagre a força vinculativa dos precedentes[32], a designação de certos casos como "casos típicos" pelo Supremo Tribunal Popular exerce influência significativa sobre a jurisprudência inferior e sobre a quantificação das indemnizações.

- Todavia, para que este sistema de três eixos se converta numa protecção efectiva da autodeterminação sexual, subsistem lacunas estruturais. Em primeiro lugar, a incriminação penal é imprecisa, baseando-se em interpretações extensivas de tipos tradicionais – como difamação (artigo 246.º), violação de informação pessoal (artigo 253.º-1) ou disseminação de material obsceno (artigo 364.º) – que não correspondem inteiramente ao bem jurídico central das **deepfakes** sexuais, isto é, a criação e difusão não consentida de representações sexuais simuladas. Em segundo lugar, a carga probatória civil é desproporcionada: as vítimas devem comprovar simultaneamente identificabilidade, verosimilhança, amplitude da difusão e dano moral, sendo recomendável que a interpretação judicial introduza presunções de causalidade e parâmetros razoáveis de compensação. Em terceiro lugar, a cooperação transfronteiriça e a respon-

[30] *Regulamento sobre a Administração dos Serviços de Síntese Profunda na Internet*, emitido pela Administração do Ciberespaço da China (Ordem n.º 1), em vigor desde 10 de Janeiro de 2023, artigos 4.º e 12.º.

[31] *Medidas Provisórias para a Administração dos Serviços de Inteligência Artificial Generativa*, emitidas conjuntamente pela Administração do Ciberespaço da China e outros sete departamentos (Ordem n.º 15), em vigor desde 15 de Agosto de 2023.

[32] Supremo Tribunal Popular, *Orientações sobre o Reforço da Proteção Judicial dos Direitos de Personalidade* (Maio de 2020), bem como os casos típicos divulgados pelos tribunais da Internet de Pequim e outros, envolvendo imitação vocal e substituição facial com recurso a inteligência artificial.

sabilização das plataformas continuam frágeis, especialmente quanto ao controlo das fontes primárias e à re-difusão internacional de conteúdos.

- Em síntese, o modelo continental chinês já assenta em três pilares: o **Código Civil**, que estabelece a base de direitos; a **Lei de Protecção de Informação Pessoal**, que define o limiar de conformidade; e o regime de síntese profunda e inteligência artificial generativa, que consagra a responsabilidade das plataformas. O passo decisivo seguinte será explicitar a autodeterminação sexual como bem jurídico protegido, preencher a lacuna penal por via legislativa ou interpretativa, aperfeiçoar as regras de prova e os modelos de indemnização e exigir mecanismos de identificação técnica e registo de algoritmos que transformem o controlo das plataformas de um modelo reactivo para um sistema preventivo, de rastreio contínuo e de responsabilização integral – assegurando, assim, uma resposta penal e civil coerente à erosão sistémica da autodeterminação sexual na era digital.

5.2. *Hong Kong – Experiência legislativa e desafios locais*

- Na Região Administrativa Especial de Hong Kong, os desafios colocados pelas tecnologias *deepfake* à autodeterminação sexual têm vindo a emergir como uma questão de relevância pública. Contudo, o ordenamento jurídico vigente ainda não cobre integralmente este novo tipo de agressão. A autodeterminação sexual, enquanto bem jurídico central do Direito Penal moderno, assenta no consentimento e no poder de cada pessoa sobre o seu próprio corpo e sobre a forma como a sua imagem sexual é representada. As imagens *deepfake*, criadas sem autorização, combinam traços faciais reais com material pornográfico, gerando representações fictícias, mas altamente verosímeis, difundidas rapidamente através da Internet. Tal prática lesa a reputação da vítima, priva-a do direito de dispor sobre a sua identidade sexual e causa danos sérios à dignidade e à saúde psicológica[33].

[33] Hong Kong SAR, *Crimes Ordinance* (Cap. 200) – Alterações introduzindo novos crimes relacionados com a privacidade sexual (*sexually intimate privacy offences*). Ver *Comunicado de Imprensa do Governo da RAEHK* / explicação legislativa no portal CLIC: https://www.info.gov.hk/gia/general/202110/08/P2021100700545p.htm.

- Em 2021, Hong Kong reformou a *Crimes Ordinance*[34], introduzindo o artigo 159AAB (voyeurismo) e os artigos 159AAC[35] a 159AAE, criminalizando a observação ou gravação não consentida de partes íntimas ou actos de natureza privada, com penas até cinco anos de prisão[36]. Estas disposições constituem uma resposta institucional aos fenómenos de gravações ilícitas e de "pornografia de vingança". Todavia, a definição de "imagem íntima" mantém-se limitada a fotografias ou vídeos reais, deixando em aberto se as imagens artificiais geradas por inteligência artificial podem ser abrangidas. Tal lacuna gera incerteza jurídica, podendo, sob uma interpretação estrita, excluir as *deepfakes* da aplicação penal.

- O artigo 161.º da mesma lei ("acesso a computador com intenção criminosa ou desonesta") poderia, em teoria, abranger a produção ou difusão de *deepfakes*, quando exista intenção de insultar, ameaçar ou obter lucro[37]. Contudo, o bem jurídico protegido é a segurança informática, não a autodeterminação sexual, o que limita substancialmente a sua aplicabilidade. No plano civil, a *Defamation Ordinance* oferece uma via de reparação, permitindo às vítimas reclamar indemnização pelos danos à reputação. Porém, a prova é onerosa: é necessário demonstrar que uma pessoa razoável acreditaria que a vítima participou efectivamente em actos obscenos, e o custo e a morosidade do processo desencorajam o recurso judicial. Trata-se, além disso, de um mecanismo individual, incapaz de gerar um efeito dissuasor generalizado.

- Quanto à *Personal Data (Privacy) Ordinance*, embora imponha restrições à recolha e utilização de dados pessoais e preveja crimes como o de *doxing* (artigo 64.º)[38], não contém disposições específicas aplicáveis a *deepfakes*. Falta-lhe, em particular, um direito de apagamento semelhan-

---

[34] *Crimes Ordinance* (Cap. 200), artigo 159AAB – *Offence of voyeurism* (*crime de voyeurismo*). Disponível em *e-Legislation Hong Kong*: https://www.elegislation.gov.hk/hk/cap200.

[35] *Ibidem*, artigo 159AAD – *Publication of images originating from voyeurism or unlawful recording* (*publicação de imagens obtidas por voyeurismo ou gravação ilícita*).

[36] *Ibidem*, artigo 159AAE – *Non-consensual publication or threat of publication of intimate images* (*publicação ou ameaça de publicação de imagens íntimas sem consentimento*).

[37] *Ibidem*, artigo 161 – *Access to computer with criminal or dishonest intent* (*acesso a computador com intenção criminosa ou desonesta*).

[38] *Personal Data (Privacy) Ordinance* (Cap. 486), *e-Legislation Hong Kong*. Disponível em: https://www.elegislation.gov.hk/hk/cap486.

te ao previsto no artigo 17.º do *Regulamento Geral sobre a Protecção de Dados* da União Europeia[39]. Assim, uma vez divulgada, a imagem *deepfake* dificilmente pode ser removida por completo, perpetuando o dano através de múltiplas redistribuições em linha. A natureza permanente e replicável destes conteúdos mantém as vítimas num estado contínuo de exposição e revitimização[40].

- Estas deficiências demonstram que o direito de Hong Kong permanece ancorado num paradigma de "imagens reais", não tendo ainda colocado a autodeterminação sexual no centro da política criminal. Sob a óptica dogmática, a violação da autodeterminação não depende da veracidade da imagem, mas da ausência de consentimento sobre a sua criação e difusão. Assim, mesmo um conteúdo inteiramente fictício pode constituir uma agressão grave à liberdade sexual. A ausência de resposta penal clara a este fenómeno revela uma lacuna legislativa. Para assegurar uma protecção eficaz, torna-se necessário rever a *Crimes Ordinance* de modo a incluir as imagens *deepfake* não consentidas no conceito de "imagens íntimas" e reconhecer explicitamente a autodeterminação sexual como bem jurídico protegido.

- Paralelamente, impõe-se a revisão da *Personal Data (Privacy) Ordinance*, criando um direito legal à eliminação e remoção de conteúdos digitais ilícitos[41]. Só através desta dupla reforma – penal e de protecção de dados – será possível dotar Hong Kong de um sistema jurídico capaz de responder às novas formas de violência sexual digital e de restaurar a dignidade e a segurança das vítimas na era tecnológica.

---

[39] Secretariado do Conselho Legislativo da Região Administrativa Especial de Hong Kong, «Resposta Escrita à Pergunta LCQ10: Regulação das Actividades de 'Doxxing'», Comunicado de Imprensa, 15 de Julho de 2020). Disponível em: https://www.info.gov.hk/gia/general/202007/15/P2020071500389.htm.

[40] Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de Abril de 2016 – Regulamento Geral sobre a Proteção de Dados (RGPD), artigo 17.º (Direito ao apagamento/direito a ser esquecido), publicado no *Jornal Oficial da União Europeia*, L119/1.

[41] Comissariado para a Privacidade de Dados Pessoais de Hong Kong, «Orientações sobre o Crime de "Doxxing"», Office of the Privacy Commissioner for Personal Data, 2021. Disponível em: https://www.pcpd.org.hk/.../guidance_doxxing_e.pdf.

5.3. *Portugal – Reformas legislativas recentes e protecção da autodeterminação sexual*

- Em 2023, Portugal aprovou a **Lei n.º 26/2023**, que introduziu alterações decisivas ao *Código Penal* e à *Lei do Comércio Electrónico*, criando um tipo penal autónomo para a "divulgação não consentida de imagens íntimas". Nos termos do artigo 193.º, na redacção então revista, comete crime quem, sem o consentimento da pessoa retratada, divulgar ou auxiliar na divulgação, por meio de comunicação social, Internet ou qualquer outro canal público, imagens, fotografias ou gravações relativas à vida privada ou sexual de outrem. A infracção é punível com pena de prisão até cinco anos e integra o título dos crimes contra a "vida privada", sendo, em regra, um crime semipúblico – dependente de queixa[42].

- Todavia, o artigo 198.º do mesmo Código prevê que, quando a conduta resultar em suicídio ou morte da vítima, ou quando o Ministério Público considerar necessário agir em defesa do seu interesse, a acção penal possa ser instaurada oficiosamente. Paralelamente, a Lei n.º 26/2023 modificou ainda a *Lei do Comércio Electrónico*, impondo obrigações positivas às plataformas digitais[43]: sempre que um prestador de serviços tenha conhecimento, ou seja notificado, da existência de conteúdos susceptíveis de constituir o referido crime, deve comunicar imediatamente o facto ao Ministério Público e proceder, no prazo de 48 horas, ao bloqueio ou remoção das páginas ou ligações relevantes, sob pena de responsabilidade legal.

- Esta reforma aproxima-se, em espírito, do mecanismo norte-americano **"TAKE IT DOWN"**[44], revelando uma convergência internacional

---

[42] Lei n.º 26/2023, de 30 de maio, que altera o Código Penal, artigo 193.º – "Quem, sem consentimento, disseminar ou contribuir para a disseminação […] de imagens, fotografias ou gravações que devassem a vida privada das pessoas, designadamente a intimidade da vida familiar ou sexual […] é punido com pena de prisão até 5 anos".

[43] *Ibidem*, artigo 198.º – "Salvo no caso do artigo 193.º, quando do crime resultar suicídio ou morte da vítima, ou quando o interesse da vítima o aconselhe, o procedimento criminal pelos crimes previstos no presente capítulo depende de queixa ou de participação".

[44] Lei n.º 26/2023, que altera simultaneamente o Decreto-Lei n.º 7/2004 e introduz disposições sobre a obrigação de notificação e remoção por parte das plataformas, determinando que os prestadores de serviços devem agir após terem conhecimento ou recebido notificação de conteúdos ilícitos, devendo proceder ao bloqueio ou remoção desses conteúdos no prazo máximo de 48 horas.

quanto à necessidade de combater a disseminação não consentida de conteúdos íntimos. Importa sublinhar que o artigo 193.º não exige que a imagem tenha origem em gravação real: basta que o conteúdo permita identificar claramente a vítima e diga respeito à sua esfera privada ou sexual. Assim, um vídeo pornográfico *deepfake* que reproduza digitalmente o rosto de uma pessoa reconhecível, divulgado sem o seu consentimento, pode enquadrar-se, por via interpretativa, neste tipo legal.

- Apesar disso, a prática judiciária enfrenta ainda desafios relevantes. Entre eles, a possibilidade de o arguido invocar que o vídeo é ficcional ou de natureza satírica[45], suscitando tensões entre liberdade de expressão e protecção da privacidade; e a dificuldade probatória associada à análise técnica das imagens, essencial para determinar o modo de geração do conteúdo e confirmar a identificabilidade da vítima.

- Com esta reforma, Portugal passou a dispor de um regime penal completo destinado a reprimir a divulgação não consentida de imagens íntimas, reforçando de forma substancial a tutela da dignidade pessoal e da autodeterminação sexual. Por contraste, o *Código Penal de Macau* ainda não contempla um tipo penal específico equivalente, limitando-se a abordar condutas próximas através dos crimes de difamação, gravação ilícita ou divulgação indevida de comunicações. Tal solução é manifestamente insuficiente para abranger a pornografia *deepfake* e outras formas de manipulação digital.

- A experiência portuguesa demonstra, assim, que Macau deverá, a médio prazo, considerar a criação de um crime autónomo de "divulgação não consentida de conteúdos sexuais", o reforço da responsabilidade das plataformas electrónicas e a atribuição ao Ministério Público de poderes de iniciativa processual em casos de relevante interesse público – medidas necessárias para colmatar as actuais lacunas normativas e alinhar

[45] De acordo com a redação atual do artigo 193.º do Código Penal (versão vigente), a norma não exige expressamente que a imagem em causa tenha sido captada de forma real, bastando que a sua divulgação ocorra "sem consentimento" e envolva "imagens, fotografias ou gravações que devassem a vida privada [...] ou a intimidade da vida familiar ou sexual". Assim, entende-se que imagens sintéticas, mas dotadas de reconhecibilidade pessoal, podem ser abrangidas pelo âmbito de proteção da disposição.

o sistema jurídico local com as tendências internacionais de protecção penal da autodeterminação sexual[46].

5.4. *Estados Unidos – Da iniciativa estadual à regulação federal*

- Nos Estados Unidos, a regulação da pornografia *deepfake* evoluiu de forma gradual, passando de iniciativas isoladas a um enquadramento federal. Inicialmente, a ausência de legislação nacional deixava as vítimas sem recurso legal. Um caso emblemático foi o da influenciadora digital **QTCinderella**, que descobriu centenas de vídeos pornográficos *deepfake* criados sem o seu consentimento, sem que houvesse base jurídica para responsabilizar os autores.

- Entre 2019 e 2020, alguns Estados começaram a legislar de forma autónoma. A Califórnia e o Texas aprovaram normas proibindo o uso de *deepfakes* para manipular processos eleitorais[47]. No mesmo ano, a Califórnia introduziu a *Assembly Bill 602*, acrescentando o artigo 1708.86.º ao *Código Civil*, que prevê responsabilidade civil para quem crie ou divulgue, sem consentimento escrito, imagens realistas de nudez ou actos sexuais geradas por computador ou inteligência artificial. O diploma permite às vítimas requerer medidas cautelares, indemnizações

[46] Acórdão da Relação de Coimbra, Apelação n.º 302/19.2T8MGL.C1, Tribunal Judicial da Comarca de Viseu – Juízo de Competência Genérica de Mangualde, Relatora: Sílvia Pires, 29 de Junho de 2021 – Decisão relativa ao *direito à imagem e aos direitos de personalidade*, na qual o tribunal reconheceu que o direito à imagem constitui um direito de personalidade protegido pelo artigo 26.º da Constituição da República Portuguesa e pelos artigos 79.º e 81.º do Código Civil, exigindo o consentimento do titular para a sua utilização. O acórdão faz igualmente referência ao Regulamento Geral sobre a Proteção de Dados (RGPD) e à Lei n.º 58/2019 como instrumentos complementares de reforço dessa proteção.

[47] *Princeton Law Journal*, «Análise sobre os riscos jurídicos das imagens deepfake», 2021 – O estudo observa que mais de 90% dos vídeos *deepfake* existentes correspondem a conteúdos pornográficos não consensuais, tendo as mulheres como principais vítimas. O relatório acrescenta que, nos Estados Unidos, apenas alguns Estados dispõem de legislação limitada sobre o tema, persistindo uma lacuna a nível federal.

por danos económicos e morais[48], e ainda compensações punitivas, excluindo expressamente o argumento de que o conteúdo seria "ficcional". A Geórgia, através do **SB 337**, criminalizou a produção e divulgação de pornografia *deepfake*; a Virgínia alterou o §18.2-386.2 do seu código penal para proibir a difusão de imagens manipuladas que representem pessoas nuas; e o Estado de Nova Iorque consagrou o direito de as vítimas intentarem acções civis contra os responsáveis pela publicação.

- Apesar desses avanços, a maioria dos Estados continuava sem legislação específica, gerando um panorama fragmentado. A protecção da liberdade de expressão ao abrigo da **Primeira Emenda** suscitava preocupações de censura excessiva. No entanto, a doutrina jurídica sustentou que a pornografia *deepfake* não consentida se enquadra nas categorias de discurso não protegido – como difamação ou obscenidade – e, portanto, pode ser legitimamente restringida.

- O consenso social crescente conduziu, em **maio de 2025**, à aprovação federal da **Lei de Remoção de Imagens Íntimas Não Consentidas** (*Non-Consensual Intimate Image Removal Act*)[49], assinada pelo Presidente, que pela primeira vez criminaliza, a nível nacional, a divulgação de conteúdos íntimos sem consentimento[50]. A lei define "falsificação digital" como qualquer criação ou modificação, por software, *machine learning* ou inteligência artificial, de imagens íntimas indistinguíveis de registos reais[51]. Estabelece que a divulgação só é lícita quando exista consentimento expresso da pessoa retratada, quando o conteúdo tenha sido

---

[48] *California AB 602*, *Bill Text: CA AB602 (2019)* – Acto legislativo aprovado pela Assembleia da Califórnia, que introduziu o *California Civil Code § 1708.86*, estabelecendo que a criação e a difusão de imagens sexuais digitais ou sintéticas sem consentimento constituem violação civil, prevendo medidas de reparação como injunções, indemnizações e danos punitivos, bem como as excepções aplicáveis. Disponível em *Legiscan, Bill Text: CA AB602*.

[49] «President Donald J. Trump Signed S. 146 into Law» – Comunicado oficial da Casa Branca informando que o Presidente sancionou o diploma em 19 de Maio de 2025.

[50] Instituto para a Indústria da Informação (*Institute for Information Industry*), Instituto de Direito e Tecnologia – Análise especializada sobre o *Take It Down Act* aprovado nos Estados Unidos em 2025, o qual, pela primeira vez a nível federal, abrange "imagens íntimas não consensuais" e "imagens *deepfake*", impondo às plataformas a obrigação de remoção no prazo de 48 horas, sob pena de sanções pela Federal Trade Commission (III – IDT, 2025).

[51] S.146 – *Take It Down Act*, que altera o *Communications Act* (47 U.S.C. § 223(h)), definindo "digital forgery" e "non-consensual intimate image", e estabelecendo as respectivas sanções.

voluntariamente tornado público ou esteja relacionado com questões de interesse público. A violação dessas condições é punível com até dois anos de prisão nos casos envolvendo adultos, e até três anos se envolver menores.

- O diploma tipifica ainda como crime o uso de tais conteúdos para ameaça ou extorsão, prevendo o confisco dos equipamentos e lucros ilícitos, bem como a obrigação de indemnização. Exige também que as principais plataformas digitais mantenham um **mecanismo rápido de remoção**, devendo eliminar o material e todas as cópias derivadas no prazo máximo de 48 horas após notificação, sob pena de sanções aplicadas pela **Federal Trade Commission**[52].

- Esta lei combina instrumentos penais e administrativos, representando um marco na luta contra a violência sexual digital e os riscos associados à inteligência artificial[53]. Embora a sua execução enfrente desafios práticos – como a escassez de peritos, a verificação técnica da prova e o risco de abuso do sistema de remoção –, a sua aprovação bipartidária reflecte o consenso político e social em torno da protecção da privacidade e da prevenção do uso abusivo da tecnologia[54]. O modelo norte-americano, que conjuga repressão penal com co-regulação das plataformas, constitui hoje uma referência internacional para a governança das *deepfakes* e para a salvaguarda da autodeterminação sexual na era digital[55].

---

[52] S.146 – *Take It Down Act* – O diploma estabelece expressamente a obrigação das plataformas de remover, no prazo máximo de 48 horas após a recepção de uma notificação, qualquer imagem íntima e as suas cópias derivadas. A lei confere ainda competência à Federal Trade Commission (FTC) para a supervisão e aplicação das sanções correspondentes.

[53] S.146 – *Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks Act* – abreviadamente designada por *Take It Down Act*, conforme consta no próprio texto legislativo.

[54] O registo do processo legislativo do S.146 no portal Congress.gov mostra que a lei foi aprovada por ampla maioria, constituindo um caso de apoio bipartidário.

[55] S.146 – *TAKE IT DOWN Act – Full Text & Summary*, consultado em Maio de 2025, disponível em: https://www.congress.gov/bill/119th-congress/senate-bill/146/text/es. SIIA, «SIIA Supports Passage of S-146, the TAKE IT DOWN Act», Comunicado de Imprensa da SIIA, 2025, disponível em: https://www.siia.net/siia-supports-passage-of-s-146-the-take-it-down-act.

5.5. *Alemanha – Direitos da personalidade e propostas de reforma penal*

Na Alemanha, embora ainda não exista um tipo penal autónomo especificamente dedicado à pornografia *deepfake*, o ordenamento jurídico vigente oferece instrumentos relevantes de protecção. A Lei Fundamental Alemã (*Grundgesetz*), através da jurisprudência consolidada do Tribunal Constitucional Federal (*Bundesverfassungsgericht*), desenvolveu o princípio do direito geral da personalidade (*allgemeines Persönlichkeitsrecht*), que abrange o direito à imagem e à reputação, proibindo a utilização ou difusão da imagem de uma pessoa sem o seu consentimento[56]. O §22 da Lei sobre o Direito à Imagem (*Kunsturhebergesetz – KUG*) estabelece que a divulgação pública de fotografias ou representações visuais de uma pessoa depende da sua autorização, salvo nas excepções previstas para reportagens de interesse público ou fins equivalentes[57].

O §201a do Código Penal Alemão (*Strafgesetzbuch – StGB*) pune a captação ou divulgação não consentida de imagens obtidas em locais privados ou protegidos, bem como a difusão de conteúdos visuais susceptíveis de causar dano grave à reputação de outrem[58]. Embora esta norma tenha sido concebida para situações envolvendo imagens reais, a prática forense demonstra que, se uma imagem *deepfake* gerada por inteligência artificial permitir a identificação inequívoca da pessoa retratada e afectar a sua honra ou privacidade, o Ministério Público pode, em certos casos, proceder à acusação com base neste dispositivo.

Não obstante, subsistem dificuldades dogmáticas e probatórias: como as *deepfakes* não resultam de uma captação física, a sua subsunção ao §201a é discutível, e a demonstração da autoria e da manipulação técnica é frequentemente complexa. Acresce que tais casos suscitam tensões entre a tutela da personalidade e as liberdades de expressão e de criação artística.

Para colmatar estas lacunas, o Governo Federal apresentou um projecto de lei que prevê a introdução do novo §201b no *StGB*, destinado a incriminar a

---

[56] «BVerfG, Urteil des Ersten Senats vom 5. Juni 1973 – 1 BvR 536/72 (*Lebach-Urteil*)», publicado em *BVerfGE* 35, pp. 202-245 – O acórdão estabeleceu um quadro de ponderação entre os direitos fundamentais consagrados no artigo 2.º, n.º 1, em conjugação com o artigo 1.º, n.º 1 (liberdade geral de personalidade e dignidade humana), e o artigo 5.º, n.º 1 (liberdade de expressão e de imprensa), exercendo profunda influência na interpretação dos direitos de personalidade no sistema constitucional alemão.

[57] Alemanha, *Kunsturhebergesetz* (KunstUrhG), artigo 22.º – Princípio do consentimento relativo ao direito de imagem: "Bildnisse dürfen nur mit Einwilligung des Abgebildeten verbreitet oder öffentlich zur Schau gestellt werden […] Nach dem Tode des Abgebildeten […] bis zum Ablaufe von 10 Jahren der Einwilligung der Angehörigen". Disponível em: https://www.gesetze-im-internet.de/kunsturhg/__22.html.

[58] Alemanha, Código Penal (*Strafgesetzbuch*, *StGB*), artigo 201a – *Verletzung des höchstpersönlichen Lebensbereichs durch Bildaufnahmen* (crime de violação da esfera de vida estritamente privada através de gravações ou imagens). Fonte oficial: *Gesetze im Internet*, disponível em: https://www.gesetze-im-internet.de/stgb/__201a.html.

criação, manipulação ou divulgação, por meios informáticos ou de inteligência artificial, de imagens indistinguíveis de conteúdos reais e que causem lesão aos direitos da personalidade de outrem, com pena máxima de cinco anos nos casos mais graves. O texto propõe ainda excepções proporcionais para reportagens jornalísticas e obras artísticas legítimas, de modo a preservar o equilíbrio entre protecção e liberdade.

De facto, a Diretiva (UE) 2024/1385 do Parlamento Europeu e do Conselho, relativa ao combate à violência contra as mulheres e à violência doméstica, inaugura uma tendência harmonizadora a nível europeu, impondo aos Estados-Membros a criminalização das *deepfakes* sexuais. Macau, embora fora da União, pode inspirar-se nesta metodologia de tipificação funcionalmente aberta e tecnologicamente neutra. Em paralelo, a referida Diretiva impôs aos Estados-Membros a obrigação de criminalizar, até 2027, a produção, manipulação e divulgação não consentidas de imagens íntimas, incluindo as geradas por inteligência artificial. Como Estado-Membro, a Alemanha deverá adaptar a sua legislação penal a estas exigências[59], sendo previsível que venha a incorporar expressamente as *deepfakes* sexuais no seu sistema penal – quer mediante a criação de um tipo autónomo, quer pela clarificação do âmbito de aplicação das normas existentes –, reforçando assim a protecção da dignidade e da autodeterminação sexual face às novas formas de violação digital.

6. Tendências internacionais e lições comparadas

Nesta senda, torna-se útil observar as tendências legislativas de outras jurisdições que enfrentam desafios semelhantes, de modo a extrair parâmetros comparativos para a realidade de Macau.

Para além dos países e regiões acima mencionados, o Reino Unido adoptou igualmente uma posição mais assertiva face à disseminação não consentida de conteúdos íntimos. O Governo britânico anunciou a inclusão desse comportamento como "priority offence"[60] ao abrigo do *Online Safety Act*, impondo às plataformas digitais a obrigação activa de detectar e remover tais conteúdos de forma célere. Em 2025, o Executivo declarou ainda a intenção de tipificar a criação de imagens *deepfake* de natureza sexual como crime autónomo, punível

[59] Directiva (UE) 2024/1385 do Parlamento Europeu e do Conselho, de 14 de Maio de 2024, relativa ao combate à violência contra as mulheres e à violência doméstica. Publicada no *Jornal Oficial da União Europeia*, disponível em *EUR-Lex*: https://eur-lex.europa.eu/eli/dir/2024/1385/oj.

[60] UK Government, «Crackdown on intimate image abuse as government strengthens online safety laws» (13 de Setembro de 2024). Disponível em: https://www.gov.uk/government/news/crackdown-on-intimate-image-abuse-as-government-strengthens-online-safety-laws.

com pena de prisão até dois anos, reforçando a resposta penal às novas formas de violência sexual digital[61].

Na Região de Taiwan, o Ministério da Justiça publicou um projecto de lei que acrescenta ao Código Penal o novo Capítulo 28-1, sob a epígrafe "Crimes de violação da privacidade sexual e de imagens falsas de carácter sexual". O tipo principal, designado "crime de criação ou divulgação de imagens falsas de carácter sexual de outrem", prevê penas até cinco anos de prisão[62], podendo atingir sete anos quando exista finalidade lucrativa[63]. Paralelamente, a Procuradoria-Geral de Taiwan esclareceu, em parecer oficial, que a produção e difusão de imagens manipuladas digitalmente – mediante técnicas de montagem ou substituição facial – constitui infracção ao artigo 319.º-4 do Código Penal, punível com prisão até cinco anos ou pena de detenção[64].

Tanto o modelo britânico como o taiwanês demonstram uma tendência convergente: o reconhecimento da necessidade de criminalizar expressamente as representações sexuais artificiais criadas sem consentimento, associando a responsabilidade penal das plataformas à protecção da dignidade e da autodeterminação sexual das vítimas.

Estas tendências revelam que, em diferentes partes do mundo, cresce o reconhecimento de que as tecnologias *deepfake* representam uma ameaça directa à autodeterminação sexual – bem jurídico essencial da liberdade e dignidade sexual –, bem como à igualdade de género e à confiança social. Os diversos ordenamentos jurídicos começam, assim, a recorrer ao instrumento legislativo para colmatar as lacunas existentes e estabelecer uma resposta normativa efectiva a este novo tipo de violência digital.

---

[61] UK Government, «Government crackdown on explicit deepfakes» (Ministry of Justice, Comunicado de Imprensa, 7 de Janeiro de 2025). Disponível em: https://www.gov.uk/government/news/government-crackdown-on-explicit-deepfakes.

[62] Ministério da Justiça da República da China (Taiwan), Comunicado de Imprensa sobre o Projeto de Alteração Parcial do Código Penal – Propõe a criação do capítulo "Crimes de Violação da Privacidade Sexual e de Imagens Sexuais Falsas", criminalizando a captação ou divulgação de imagens sexuais, bem como de *deepfakes*, sem consentimento, com pena máxima de sete anos de prisão (Ministério da Justiça, 2022).

[63] Conselho Executivo da República da China (Taiwan)/Ministério da Justiça, «Crimes de Deepfake e outras condutas ilícitas aprovados pelo Conselho Executivo – Alteração Parcial do Código Penal da República da China». Disponível em: https://www.ey.gov.tw/Page/9277F759E41CCD91/afd0f04a-1de4-4c97-a874-5173cd34d43f.

[64] Procuradoria-Geral de Segunda Instância de Taiwan, «Responsabilidade Penal Relacionada com Imagens Sexuais Digitais», sítio oficial do Ministério Público (secção "Descrição dos Casos"). Disponível em: https://www.tph.moj.gov.tw/4421/4475/632364/1310236/post [consultado em 30/09/2025].

7. Propostas de reforma

Perante os desafios que as tecnologias *deepfake* colocam à justiça penal, Macau deve actuar simultaneamente nos domínios do direito substantivo e do mecanismo de aplicação da lei, a fim de reforçar a protecção jurídica e a eficácia da actuação criminal.

- Criar um tipo legal autónomo que permita criminalizar expressamente as condutas associadas à pornografia *deepfake*. Com base nas experiências internacionais, recomenda-se que o Código Penal de Macau inclua um novo artigo no capítulo dos crimes sexuais, punindo a produção, difusão, transmissão, venda ou exibição pública, sem consentimento, de imagens virtuais de nudez ou de actos sexuais geradas por meios informáticos ou outras tecnologias de síntese. O tipo penal deve abranger diversas modalidades, como vídeos, imagens ou gravações de áudio que simulem a participação da vítima em actos de natureza sexual. Na elaboração legislativa, importa deixar claro que o bem jurídico primordialmente protegido é a autodeterminação sexual, enquanto a privacidade e a reputação são bens jurídicos complementares, cuja tutela decorre reflexamente da violação do primeiro, enfatizando que, ainda que o conteúdo não seja real, tal constitui violação da esfera sexual da vítima. À semelhança do modelo em fase de discussão em Taiwan, as condutas poderiam ser divididas em duas categorias – produção e divulgação – com escalas penais distintas e agravantes específicas para os casos de reincidência, intuito lucrativo ou difusão em larga escala. Através desta incriminação específica, colmatar-se-ia a actual lacuna normativa e afirmar-se-ia inequivocamente que qualquer utilização de tecnologias de inteligência artificial para violar a dignidade sexual de outrem constitui crime intolerável num Estado de Direito.

- Reforçar a capacidade de perícia digital e aperfeiçoar o regime da prova electrónica. O sistema judicial deve acompanhar o ritmo da evolução tecnológica, investindo recursos na melhoria da detecção e investigação de conteúdos *deepfake*. Por um lado, é necessário criar ou fortalecer laboratórios de perícia digital, dotando-os de ferramentas avançadas de detecção e de algoritmos especializados, bem como formar peritos capazes de identificar a autenticidade das gravações e rastrear a pegada digital dos seus autores – por exemplo, mediante a introdução de tecnologias de análise de imagem baseadas em inteligência artificial, capazes de detectar anomalias visuais e marcas digitais (*watermarks*). Por outro lado, no plano jurídico, devem ser elaboradas directrizes sobre a recolha, conservação e

apresentação de prova electrónica, estabelecendo padrões específicos para a admissão de vídeos *deepfake* como elementos probatórios. Tais normas deveriam exigir que as autoridades conservassem o endereço original, as características digitais e um relatório técnico de verificação de autenticidade emitido por perito qualificado, reforçando a credibilidade da prova em tribunal. Se necessário, o Código de Processo Penal poderia ser revisto para introduzir um regime de perito técnico em prova electrónica, permitindo que especialistas certificados prestem pareceres sobre tecnologias *deepfake* e auxiliem os juízes na apreciação de provas altamente técnicas.

- Criar um mecanismo interdepartamental de coordenação dedicado ao combate à violência sexual digital. A pornografia *deepfake* situa-se frequentemente na intersecção entre o cibercrime e a criminalidade sexual, exigindo uma resposta integrada. Recomenda-se que Macau estabeleça um grupo de trabalho interdepartamental, ou designe uma unidade especializada, por exemplo, no seio da Polícia Judiciária, reunindo agentes das divisões de criminalidade informática e de investigação criminal, com competência específica para investigar a difusão de conteúdos de natureza sexual ilícita na Internet. Simultaneamente, deve ser reforçada a articulação entre o Ministério Público, a Polícia e os tribunais, mediante a criação de protocolos operacionais que abranjam a partilha de informação, a definição de prioridades de investigação, a transferência e preservação de provas. No plano da assistência às vítimas, é essencial assegurar acompanhamento após a denúncia, disponibilizando apoio técnico para localizar e solicitar a remoção de vídeos *deepfake* divulgados em rede, prevenindo assim danos secundários. A institucionalização deste mecanismo de coordenação garantirá uma actuação contínua e coerente – desde a denúncia até ao julgamento –, melhorando a eficiência processual e a capacidade do sistema penal em responder de forma célere e eficaz a este novo tipo de violência digital.

- Reforçar a cooperação judiciária transfronteiriça e a governação das plataformas digitais. A difusão de conteúdos pornográficos *deepfake* apresenta um carácter intrinsecamente transnacional, exigindo que Macau amplie activamente os seus canais de cooperação internacional. Por um lado, deve recorrer aos mecanismos de colaboração regional para partilhar experiências e informações com a China Continental e com os países de língua portuguesa, aprendendo com os modelos eficazes adoptados noutras jurisdições. Por outro lado, Macau poderá integrar

ou promover a criação de uma rede internacional de combate à violência sexual em linha, visando facilitar a troca de provas, a extradição de suspeitos e a assistência mútua em matéria penal. Relativamente às principais plataformas e redes sociais, impõe-se um reforço da comunicação e da cooperação. Apesar da pequena dimensão do território, Macau pode recorrer a organizações regionais ou à mediação do Governo Central para negociar com empresas tecnológicas internacionais, de modo a assegurar respostas rápidas à difusão de conteúdos *deepfake* – por exemplo, seguindo o modelo norte-americano que exige a remoção em 48 horas. Seria igualmente recomendável criar uma base legal local que permita aos tribunais ordenar aos prestadores de serviços de Internet a remoção imediata de conteúdos ilícitos e a conservação dos respectivos dados para efeitos de investigação. Num contexto global de combate aos riscos associados às *deepfakes*, Macau deve alinhar-se com os padrões internacionais e intensificar a cooperação técnica e jurídica, reduzindo assim as suas vulnerabilidades em matéria de jurisdição e de capacidade tecnológica.

8. Conclusão

O desenvolvimento tecnológico coloca ao Direito desafios sem precedentes, mas também o impele a evoluir continuamente para salvaguardar os direitos fundamentais dos cidadãos. As deepfakes representam uma nova forma de criminalidade digital que viola directamente a autodeterminação sexual – o núcleo essencial da liberdade e dignidade sexual – exigindo uma resposta dogmática e institucional adequada por parte do Direito Penal.

Conclui-se, portanto, que o ordenamento penal macaense evidencia uma lacuna estrutural de tipicidade, na medida em que o uso não consentido da imagem de outrem para fins libidinosos ou pornográficos não constitui, à presente data, objecto de proibição penal expressa. Nem o capítulo dos crimes contra a autodeterminação sexual, nem o dos crimes contra a honra se revelam aptos a abranger adequadamente esta nova modalidade de agressão à liberdade e à dignidade sexual. Tal constatação reforça a necessidade de uma resposta legislativa autónoma, tecnicamente precisa e tecnologicamente neutra, capaz de garantir a coerência sistemática do Direito Penal perante os novos fenómenos de simulação e manipulação digital.

Embora Macau constitua uma jurisdição de pequena dimensão, não está imune aos riscos e impactos de uma tendência global. Impõe-se, por isso, uma actuação preventiva: criar tipos legais específicos que preencham as lacunas existentes, fortalecer a capacidade pericial e a cooperação transfronteiriça, e conso-

lidar mecanismos de coordenação interinstitucional e internacional que assegurem a eficácia da repressão penal.

Em última análise, a tutela penal da autodeterminação sexual na era digital requer uma dogmática dinâmica e sensível às novas formas de agressão simbólica e virtual. A incriminação das *deepfakes* sexuais não constitui apenas uma opção de política criminal, mas antes uma exigência de dignidade humana, reafirmando o papel do Direito Penal enquanto instrumento contemporâneo de protecção efectiva dos direitos fundamentais.

NEURODIREITOS NA ERA DIGITAL: O CONFLITO ENTRE INTELIGÊNCIA ARTIFICIAL E LIBERDADE MENTAL

*Pedro Henrique Engler Urso**
*Matheus Raphael Pova Marinho Dantas***

Resumo: Este artigo analisa a parceria entre o governo de Donald Trump e as grandes empresas de tecnologia sob a perspectiva da teoria do totalitarismo de Hannah Arendt e dos debates contemporâneos sobre neurodireitos. Partindo da noção arendtiana de dominação total, caracterizada pela manipulação sistemática das consciências e pela eliminação da espontaneidade humana, investiga-se como o uso político das plataformas digitais, potencializado pelo avanço das inteligências artificiais, pode gerar novas modalidades de controle social. Frente a esse cenário, argumenta-se pela urgência de estabelecer proteções jurídicas específicas para resguardar a autonomia cognitiva e a integridade mental dos indivíduos em sociedades cada vez mais conectadas.

* Pós-Graduado em Direito da União Europeia pela Universidade de Coimbra, Graduando em direito pela Universidade Presbiteriana Mackenzie, Graduado em Comércio Exterior. E-mail: *pedroheurso17@gmail.com.*

** Pós-graduando em Gestão de Vendas pela Universidade de São Paulo, Graduando em Relações Internacionais pela Universidade Federal de São Paulo, Graduado em Comércio Exterior. E-mail: *matheusdantas1705@gmail.com.*

Palavras-chave: Inteligência Artificial; Neurodireitos; *Big Techs*; Hannah Arendt; Donald Trump.

Abstract: This article analyzes the partnership between the Donald Trump administration and large technology companies from the perspective of Hannah Arendt's theory of totalitarianism and contemporary debates on neurorights. Based on Arendt's notion of total domination, characterized by the systematic manipulation of consciousness and the elimination of human spontaneity, the article investigates how the political use of digital platforms, enhanced by the advancement of artificial intelligence, can generate new forms of social control. Given this scenario, the article argues for the urgent need to establish specific legal protections to safeguard the cognitive autonomy and mental integrity of individuals in increasingly connected societies.

Keywords: Artificial intelligence; Neurorights; *Big Techs*; Hannah Arendt; Donald Trump.

Introdução

A era digital deu início a uma transformação discreta, mas profunda, na maneira como as sociedades produzem conhecimento, organizam o debate público e tomam decisões coletivas. Historicamente, a formação da opinião social estava ligada a instituições tradicionais como partidos políticos, universidades e meios de comunicação convencionais. Hoje, porém, ela é largamente intermediada por plataformas digitais que funcionam com algoritmos opacos, desenvolvidos e geridos por grandes corporações privadas. Nesse novo cenário informacional, a lógica da programação comportamental tende a substituir o espaço da deliberação racional, remodelando simultaneamente a estrutura da cognição humana e a própria experiência democrática.

Este estudo examina os impactos dessa transformação tecnológica sobre a autonomia mental dos indivíduos e os perigos que ela representa para as democracias atuais. A partir de casos paradigmáticos, como o escândalo Cambridge Analytica, são analisadas as táticas de modulação emocional, engenharia comportamental e direcionamento de conteúdos possibilitadas por sistemas algorítmicos. Além disso, ressalta-se que o avanço das inteligências artificiais, cada vez mais presentes na mediação da informação, acrescenta uma camada adicional de complexidade: seus vieses inerentes, fruto de dados históricos e decisões humanas anteriores, podem intensificar desigualdades, reforçar estereótipos sociais e distorcer ainda mais os processos de formação da opinião pública. Assim, essas tecnologias não se limitam a transmitir informação, mas atuam diretamente, moldando percepções e restringindo o exercício de uma cidadania crítica.

Nesse contexto, discute-se a emergência dos neurodireitos como resposta jurídica e ética às ameaças que afetam diretamente a mente humana. Mais do que um conjunto de garantias individuais, esses direitos apresentam uma nova estrutura jurídica destinada a proteger a integridade cognitiva dos indivíduos frente à crescente invasão das tecnologias digitais. Por fim, propõe-se uma reflexão sobre a necessidade urgente de uma regulação sólida e interdisciplinar, capaz de preservar a liberdade mental enquanto bem público fundamental. Sem essa proteção, há o risco de que a democracia se transforme em um simulacro, capturada por arquiteturas invisíveis de vigilância, manipulação e controle.

1. Mentes programadas: cognição humana na era algorítmica

Vivemos uma era em que o espaço público foi profundamente transformado. Enquanto no passado o debate político e a formação da opinião pública ocorriam por meio de instituições tradicionais, como partidos, sindicatos, universidades e a imprensa convencional, hoje, esses processos se dão majoritariamente em plataformas digitais controladas por empresas privadas. Essas companhias operam algoritmos opacos que organizam e filtram o fluxo informacional direcionado a bilhões de pessoas, assumindo um papel central, ainda que discreto, na construção da realidade social. Nesse contexto, as *Big Techs* conquistam um protagonismo sem precedentes, escapando dos mecanismos convencionais de regulação e responsabilidade democrática.

Observa-se, assim, o surgimento de uma nova modalidade de populismo, que não se apoia mais em líderes carismáticos ou discursos inflamados, mas no trabalho de engenheiros de dados, estrategistas digitais e especialistas em propaganda computacional. Esses profissionais da manipulação, conhecidos como "engenheiros do caos", utilizam ferramentas digitais avançadas para moldar emoções, direcionar percepções e alimentar indignações. A lógica que os orienta é a da desinformação, da polarização e da segmentação radical dos públicos, substituindo o debate de ideias por narrativas calibradas para provocar engajamento emocional e, consequentemente, controlar comportamentos em grande escala[1].

O escândalo da Cambridge Analytica, revelado em 2018, expôs os riscos concretos desse novo paradigma. A empresa coletou dados pessoais de milhões de usuários do *Facebook* sem consentimento adequado e os empregou para construir perfis psicológicos extremamente detalhados. Esses dados foram usados em campanhas políticas altamente segmentadas, com impacto significativo tanto no referendo do Brexit quanto na eleição de Donald Trump para a presidência dos Estados Unidos. A premissa era simples, porém devastadora: identificar vulnerabilidades cognitivas para manipular decisões eleitorais em massa[2].

Esses episódios revelam a fragilidade estrutural da democracia diante de tecnologias projetadas para capturar atenção, maximizar o engajamento e modular comportamentos. Ao privilegiar conteúdos emocionalmente intensos, frequentemente sensacionalistas, falsos ou polarizadores, os algoritmos reforçam

---

[1] Giuliano Da Empoli, *Engenheiros do caos: Como a desinformação e as* fake news *estão transformando a política e a sociedade*, Tradução de Claudio Carina, 1.ª ed., Rio de Janeiro: Zahar, 2020.

[2] Carole Cadwalladr; Emma Graham-Harrison, «Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach», *The Guardian*, Londres, 17 mar. 2018, disponível em: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election [Acesso em: 29/04/2025].

crenças pré-existentes e criam bolhas informacionais que isolam os indivíduos do contraditório. Está em jogo, portanto, não apenas a qualidade do debate público, mas a própria autonomia cognitiva dos cidadãos em uma era de estímulo digital contínuo[3].

As redes sociais deixaram de ser meros canais de comunicação para se tornar instrumentos de engenharia comportamental e vigilância sistêmica. A coleta massiva de dados, combinada com sistemas de aprendizado de máquina, permite não apenas prever preferências, mas induzi-las em tempo real. O público, antes consumidor passivo de informação, torna-se objeto de experimentos constantes: cada clique, reação ou tempo de visualização serve para aprimorar estratégias de persuasão, ajustando mensagens que provocam medo, desejo, raiva ou aceitação. Nesse modelo, o poder algorítmico não mede apenas comportamentos, ele molda a própria estrutura da cognição individual e coletiva[4].

Além disso, surge um problema crescente: os vieses incorporados às inteligências artificiais amplamente utilizadas pelas plataformas digitais. Como ressalta Lago[5], sistemas de inteligência artificial (IA) treinados com dados históricos oriundos de uma sociedade desigual tendem a replicar e amplificar essas distorções. Algoritmos de recomendação podem reforçar preconceitos; modelos de classificação podem discriminar no acesso a crédito, emprego ou informação; assistentes virtuais e sistemas de análise de linguagem podem privilegiar determinados padrões linguísticos ou culturais. Esse é um problema que transcende o aspecto técnico, configurando um desafio ético e político, pois tais vieses operam silenciosamente para estruturar o que vemos, pensamos e desejamos, afetando profundamente a autodeterminação individual e a justiça social[6].

A preocupação ultrapassa, portanto, a manipulação eleitoral. Plataformas como *Facebook*, *YouTube*, *Instagram* e *TikTok* influenciam diariamente nossos hábitos, sentimentos, crenças e valores. O que consumimos, como nos informamos, como reagimos às notícias e mesmo como nos posicionamos politicamente está mediado por algoritmos invisíveis, mas extremamente poderosos. A chamada "economia da atenção" atua segundo princípios contrários à racionalidade

---

[3] Jeff Orlowski (dir.); Davis Coombe; Vickie Curtis (roteiro), *O Dilema das Redes*, Netflix, 2020, disponível em: https://www.netflix.com/pt/title/81254224 [Acesso em: 29/04/2025].

[4] Shoshana Zuboff, *A era do capitalismo de vigilância: A luta por um futuro humano na nova fronteira do poder*, Rio de Janeiro: Intrínseca, 2020.

[5] Ronaldo Assunção Sousa do Lago, «A ética, o viés e a inteligência artificial», in *Consultor Jurídico*, São Paulo, 8 jul. 2023, disponível em: https://www.conjur.com.br/2023-jul-08/ronaldo-lago-etica-vies-inteligencia-artificial/ [Acesso em: 21/05/2025].

[6] Ronaldo Assunção Sousa do Lago, «A ética, o viés e a inteligência artificial», in *Consultor Jurídico*, São Paulo, 8 jul. 2023, disponível em: https://www.conjur.com.br/2023-jul-08/ronaldo-lago-etica-vies-inteligencia-artificial/ [Acesso em: 21/05/2025].

crítica: valoriza o vício, a reatividade e a fragmentação da atenção, criando um ambiente onde a reflexão é substituída pela impulsividade.

Historicamente, este é um deslocamento radical. A esfera pública moderna, construída ao longo de séculos como espaço de deliberação racional e formação da vontade coletiva, está sendo substituída por uma arena de estímulos fragmentados, narrativas emocionalmente orientadas e verdades concorrentes. Algoritmos operam como "armas de destruição matemática", capazes de impactar profundamente a coesão social, a confiança institucional e a própria ideia de verdade compartilhada[7].

Diante dessa nova configuração, torna-se imprescindível repensar o papel do Direito. O controle sobre dados pessoais, a transparência na lógica algorítmica, a responsabilização das plataformas digitais e a proteção da integridade cognitiva figuram como pilares centrais de uma nova agenda regulatória do século XXI. Ignorar esses desafios é aceitar que a democracia se transforme em uma simulação formalmente livre, mas efetivamente capturada por mecanismos de manipulação invisível e altamente sofisticados[8].

Este cenário prepara o terreno para o debate que será desenvolvido no próximo capítulo, centrado na necessidade de garantir neurodireitos: uma nova geração de direitos fundamentais voltados à proteção da liberdade mental, da identidade pessoal e da autodeterminação cognitiva frente à crescente influência das tecnologias digitais. Quando o poder incide diretamente sobre a mente humana, proteger o pensar torna-se um ato essencial de resistência e de preservação da dignidade.

2. Neurodireitos em debate: a última fronteira da autonomia humana

Sob uma ótica evolutiva, pode-se afirmar que o Direito representa uma das ferramentas mais elaboradas criadas pela humanidade para estruturar e regulamentar a convivência social. Ao longo da história, sociedades complexas e com densidade populacional sempre dependeram de sistemas jurídicos robustos, que garantissem segurança mínima e previsibilidade nas relações interpessoais. Contudo, na era contemporânea, marcada pela hiperconectividade, automação de processos e a crescente influência de algoritmos na experiência tanto

[7] Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York: Crown Publishing, 2016.

[8] M. Ienca; R. Andorno, «Towards new human rights in the age of neuroscience and neurotechnology», *Life Sci Soc Policy* 13, 5, 2017, disponível em https://doi.org/10.1186/s40504-017-0050-1 [Acesso em: 05/05/2025].

individual quanto coletiva, surge um desafio inédito para a teoria e prática do Direito: a proteção da liberdade mental frente a tecnologias capazes de manipular cognições, emoções e comportamentos. É nesse contexto que emergem os neurodireitos[9].

Nesse contexto, o neurodireito configura-se como uma área interdisciplinar que busca entender as interseções entre as neurociências e o Direito. Inicialmente, esteve ligado principalmente à aplicação de evidências neurocientíficas em processos judiciais, aproximando-se do campo conhecido como "neurociência forense". Com o passar do tempo, sua abrangência ampliou-se, incluindo debates filosóficos, éticos, políticos e tecnológicos. O conceito aborda temas fundamentais como a liberdade individual, dignidade humana, autonomia da vontade e, sobretudo, a privacidade mental, especialmente diante da possibilidade técnica de captar, interpretar e até influenciar a atividade cerebral por meio de interfaces neurais, algoritmos sofisticados e sistemas de inteligência artificial[10].

Em 2017, começaram a surgir os primeiros marcos relevantes no debate sobre neurodireitos, com publicações que evidenciavam o despreparo do sistema jurídico internacional frente aos avanços das neurotecnologias[11]. No mesmo ano, pesquisadores da neurociência lançaram um artigo de grande impacto, considerado um ponto de disrupção no tema, que listava quatro prioridades éticas essenciais para o desenvolvimento de neurotecnologias e inteligência artificial[12]. Entre os pontos destacados estavam a proteção da privacidade mental, a preservação da identidade pessoal, a garantia de acesso equitativo às inovações e a atenção necessária aos vieses embutidos nos algoritmos. O documento serviu como alerta para a urgência de conduzir o progresso tecnológico baseado em princípios éticos sólidos[13].

Com base nessas preocupações, foi instituída uma fundação internacional dedicada à promoção e proteção dos neurodireitos, intitulada *NeuroRights*

---

[9] Carlos Marden; Leonardo Martins Wykrota, «Neurodireito: o início, o fim e o meio», *Revista Brasileira de Políticas Públicas*, Brasília, v. 8, n. 2, jul./dez. 2018, DOI: 10.5102/rbpp.v8i2.5307.

[10] Renato César Cardoso, «Neurodireito e neurociência do livre-arbítrio: uma revisão crítica e apontamentos para o Direito Penal», *Revista Internacional Consinter de Direito*, v. 8, n. 1, p. 93 (pp. 91-120), 2023. DOI: 10.46274/1809-192XRICP2023v8n1p91-120.

[11] Maria Clara Nolasco Viana; Carolina Silva Porto & Clara Cardoso Machado Jaborandy, «Neurodireitos e privacidade mental sob a ótica do episódio Engenharia Reversa em Black Mirror: proteção ou controle? Interfaces Científicas», *Direito*, Aracaju, v. 2, 2023, pp. 352-368, DOI: 10.17564/2316-381X.2023v9n2p352-368.

[12] M. Ienca; R. Andorno, «Towards new human rights in the age of neuroscience and neurotechnology», *op. cit.*

[13] Ana Maria D'Ávila Lopes, *Neurodireitos: proteção à mente humana e à privacidade*, Porto Alegre: OAB/RS, 2022, p. 21, disponível em: https://www.oabrs.org.br/arquivos/file_676440cc53531.pdf [Acesso em: 25/04/2025].

Foundation. O objetivo dessa entidade é defender a privacidade mental, a identidade pessoal e o livre arbítrio dos indivíduos. Além disso, visa assegurar que os benefícios das novas tecnologias sejam distribuídos de forma justa entre os diversos segmentos da sociedade, prevenindo que os desenvolvedores perpetuem preconceitos ou discriminações em seus sistemas. Informações detalhadas sobre as iniciativas e objetivos da fundação estão disponíveis em seu *site* oficial[14].

Diante dos desafios contemporâneos impostos pelas *Big Techs*, a discussão sobre neurodireitos surge como um tema urgente e relevante. Embora seja um conceito relativamente novo, os neurodireitos tendem a ganhar cada vez mais espaço nos debates jurídicos e sociais da atualidade. Essa crescente importância reflete o avanço acelerado das inovações tecnológicas, que não apenas permeiam nosso cotidiano, mas também transformam profundamente as formas de interação com o mundo, inclusive em nossas esferas mais íntimas e subjetivas[15].

Da mesma forma que a inteligência artificial já se estabeleceu como uma força transformadora nas esferas pessoal, profissional e empresarial, as neurotecnologias avançam rapidamente. Essa aceleração justifica a crescente preocupação de estudiosos e especialistas com os limites éticos, os possíveis impactos sociais e os mecanismos regulatórios necessários para proteger a mente humana de intervenções indevidas.

Nesse sentido, torna-se evidente que a simples existência de princípios éticos ou declarações de boas intenções por parte das empresas de tecnologia não é suficiente para garantir a proteção da liberdade mental. É urgente que os Estados e organismos internacionais avancem na regulamentação das *Big Techs*, de modo a estabelecer limites claros quanto ao uso de dados neurais, algoritmos preditivos e técnicas de modulação comportamental. As grandes plataformas digitais detêm não apenas informações sobre os hábitos dos usuários, mas passam a acessar e interferir em aspectos cada vez mais sensíveis da cognição humana, emoções, impulsos, tomadas de decisão e até preferências inconscientes. Essa capacidade de influenciar o pensamento e o comportamento coloca em risco a autonomia individual e pode, em larga escala, comprometer os próprios fundamentos da democracia, como o livre-arbítrio, o voto consciente e o debate público plural[16].

---

[14] Renato César Cardoso, «Neurodireito e neurociência do livre-arbítrio: uma revisão crítica e apontamentos para o Direito Penal», *Artigo de periódico. Universidade Federal de Minas Gerais*. Disponível em: https://repositorio.ufmg.br/items/de3746c1-f5ce-4fc1-99bf-22c13266ea15 [Acesso em: 11 nov. 2025].

[15] M. Ienca; R. Andorno, «Towards new human rights in the age of neuroscience and neurotechnology», *op. cit.*

[16] Giuliano Da Empoli, *Engenheiros do caos: Como a desinformação e as* fake news *estão transformando a política e a sociedade*, Tradução de Claudio Carina, 1.ª ed., Rio de Janeiro: Zahar, 2020.

A proposta de reconhecimento dos neurodireitos como uma categoria autônoma de direitos fundamentais decorre da constatação de que as tecnologias neurodigitais produzem riscos inéditos, não plenamente abarcados pelos direitos clássicos à liberdade e privacidade. A esfera mental, entendida como espaço último da autodeterminação humana, não se confunde integralmente com esses direitos, pois envolve dimensões específicas da cognição, da memória e da vontade que podem ser diretamente acessadas, monitoradas ou manipuladas por dispositivos neurotecnológicos. Trata-se da tutela de um novo bem jurídico, a liberdade cognitiva, a salvaguarda da mente enquanto domínio inviolável da consciência, contra interferências externas capazes de alterar percepções, crenças ou processos decisórios sem o consentimento informado do indivíduo. Trata-se de proteger não apenas o "direito de pensar livremente", mas o "direito de manter-se senhor dos próprios pensamentos, contra influências externas"[17].

Nesse sentido, a autonomização dos neurodireitos não representa uma redundância normativa, mas sim uma resposta evolutiva do ordenamento jurídico aos desafios do tempo presente. Conforme o entendimento de Norberto Bobbio[18], os direitos não surgem todos de uma vez, mas se afirmam historicamente "à medida que novos valores são reconhecidos como dignos de tutela e novas necessidades emergem na vida social". Assim como a Revolução Industrial exigiu a positivação de direitos trabalhistas e o século XX demandou a consolidação dos direitos de personalidade, o advento da era neurotecnológica impõe o reconhecimento de novos direitos voltados à proteção da mente humana. A institucionalização dos neurodireitos, portanto, não é mera inovação teórica, mas expressão da contínua expansão do catálogo de direitos fundamentais diante das transformações tecnocientíficas que redefinem o próprio conceito de humanidade[19].

Esse debate vem gradualmente deixando o campo teórico para alcançar reconhecimento jurídico concreto em diferentes ordenamentos. No Brasil, o tema já encontra expressa salvaguarda na Constituição do Estado do Rio Grande do Sul que, em 2023, incluiu a proteção da liberdade cognitiva e da integridade mental entre os direitos e garantias fundamentais, reconhecendo a necessidade de preservar o indivíduo frente a tecnologias capazes de interferir em seus processos mentais. No plano nacional, discute-se a inserção dos neurodireitos

---

[17] Henrique Segolin Molina; Michel Sachs Feuz, «Desafios jurídicos da neurociência do livre-arbítrio: salvaguarda da integridade mental e reconhecimento dos dados neurais como dados pessoais sensíveis», in José Humberto Fasano Filho (org.), *O futuro dos neurodireitos*, Amanda Cunha e Mello Smith Martins; Bruno Farage da Costa Felipe; Mariana Ruzzi (coord.), Francisco Cavalcante de Sousa (coord. institucional), 1.ª ed., [S.l.]: Legal Grounds Institute, 2025.

[18] Norberto Bobbio, *Teoria geral dos direitos*, Rio de Janeiro: Editora Campus, 1992.

[19] Norberto Bobbio, *Teoria geral dos direitos*, Rio de Janeiro: Editora Campus, 1992.

na reforma do Código Civil brasileiro, com o objetivo de atualizar o sistema jurídico para abranger direitos ligados à identidade e à privacidade mental em meio à transformação digital. Internacionalmente, o Chile tornou-se pioneiro ao constitucionalizar os neurodireitos, por meio da reforma de 2021, que consagrou a proteção da integridade física e psíquica, assegurando que o desenvolvimento científico e tecnológico respeite a liberdade e a dignidade mental das pessoas. Esses avanços demonstram que o tema, antes restrito ao campo acadêmico, começa a ser efetivamente reconhecido e incorporado à estrutura normativa contemporânea[20].

Proteger a mente tornou-se, portanto, um desafio jurídico e político de primeira ordem. A ausência de uma estrutura normativa robusta abre caminho para a consolidação de um mercado de influência cognitiva, em que indivíduos se tornam alvos de intervenções sutis, contínuas e imperceptíveis, promovidas por sistemas algorítmicos otimizados para captar atenção, moldar decisões e direcionar condutas. A regulamentação não deve ter como objetivo inibir a inovação, mas sim garantir que ela ocorra dentro de parâmetros compatíveis com a dignidade humana, a privacidade mental e os direitos fundamentais. Ao estabelecer regras claras para o uso de tecnologias neurodigitais, o Direito reafirma seu papel essencial: o de proteger o indivíduo diante dos novos poderes que surgem no século XXI.

3. O império invisível: *Big Techs*, IA e o risco do totalitarismo mental

Nos últimos anos, as democracias contemporâneas passaram por transformações estruturais profundas, não apenas no campo das disputas narrativas, mas, sobretudo, na arquitetura tecnológica que sustenta e influencia os processos políticos. A incorporação crescente de tecnologias baseadas em inteligência artificial ao ecossistema informacional e à dinâmica eleitoral tem inaugurado um novo paradigma de poder, baseado na manipulação comportamental, no controle cognitivo e na erosão silenciosa das liberdades individuais. Nesse cenário, a eleição de Donald Trump, em 2024, pode ser interpretada como um marco simbólico e funcional desse processo, evidenciando como a convergência entre algoritmos, plataformas digitais e ausência de regulação efetiva pode configurar novas formas de dominação política.

---

[20] Pedro Henrique Engler Urso; Rayana Clara Pinheiro e Fernando Ramos Evanjelho, *«Neurodireitos: análise da reforma do Código Civil de 2002 e a importância do Direito Digital»,* in José Humberto Fasano Filho (org.), *O futuro dos neurodireitos*, Amanda Cunha e Mello Smith Martins; Bruno Farage da Costa Felipe; Mariana Ruzzi (coord.), Francisco Cavalcante de Sousa (coord. institucional), 1.ª ed., [S.l.]: Legal Grounds Institute, 2025.

Trump foi eleito com 312 votos no Colégio Eleitoral e, diferentemente de 2016, também conquistou a maioria do voto popular, tornando-se o primeiro republicano em mais de duas décadas a atingir esse feito. Compreender essa vitória exige ultrapassar as variáveis clássicas da ciência política. O papel desempenhado por sistemas automatizados, algoritmos preditivos e inteligência artificial na veiculação segmentada de conteúdos moldou o comportamento eleitoral de forma sutil, porém decisiva. Como alerta Zuboff, essas tecnologias não visam apenas conhecer os usuários, mas modificar suas condutas conforme interesses estratégicos[21].

A campanha de 2024 explorou intensamente a lógica algorítmica, com destaque para o uso de IA generativa na criação de narrativas fragmentadas, vídeos altamente emocionais e mensagens polarizadoras difundidas por redes automatizadas e perfis influenciadores. Essas ferramentas foram desenhadas para ativar respostas afetivas específicas: medo, raiva, ressentimento, tornando a racionalidade política refém de uma estimulação emocional constante. Como observa O'Neil, os algoritmos não são neutros: eles codificam preconceitos, automatizam desigualdades e ampliam distorções[22].

Esse processo foi agravado após a tentativa de assassinato de Trump, em julho de 2024. O episódio foi imediatamente apropriado pelas plataformas digitais como gatilho narrativo: vídeos produzidos por IA, *memes* virais e depoimentos artificiais construíram a imagem de Trump como mártir político. A viralização desses conteúdos revelou o poder das redes sociais e de seus algoritmos na criação de realidades afetivas paralelas, indiferentes à verificação dos fatos, mas altamente eficazes na modulação do comportamento político. A vitória de Trump, portanto, não se deu apenas nas urnas, mas no terreno digital da cognição induzida algoritmicamente.

Nesse ambiente, o indivíduo deixa de ser sujeito autônomo e passa a ser matéria-prima para os modelos preditivos de IA. A coleta massiva de dados, que inclui desde hábitos de consumo até padrões emocionais, permite que as plataformas moldem a percepção da realidade e antecipem reações futuras. Trata-se do que Zuboff define como capitalismo de vigilância, no qual a mente humana se converte no novo território de extração e exploração[23]. Como reforça Harari, quem controla os dados controla os sentimentos e quem controla os sentimentos, controla as pessoas[24].

---

[21] Shoshana Zuboff, *A era do capitalismo de vigilância…*, *op. cit.*, p. 19.

[22] Cathy O'Neil, *Weapons of Math Destruction…*, *op. cit.*, p. 29.

[23] Shoshana Zuboff, *A era do capitalismo de vigilância…*, *op. cit.*, p. 328.

[24] Yuval Noah Harari, *21 lições para o século 21*, São Paulo: Companhia das Letras, 2018, p. 69.

A posse de Trump em 2025 simbolizou essa fusão entre poder político e tecnológico. Estiveram presentes os principais executivos das *Big Techs*: Mark Zuckerberg (*Meta*), Sundar Pichai (*Google*), Jeff Bezos (Amazon), Sam Altman (*OpenAI*), Tim Cook (*Apple*) e Elon Musk (*Tesla/X*), este último nomeado para liderar o recém-criado Departamento de Eficiência Governamental, responsável pela implementação de sistemas de IA na gestão pública. A nomeação não foi apenas um gesto simbólico: ela marcou a consolidação de um modelo de governança tecnocorporativa, no qual a política é mediada, dirigida e monitorada por tecnologias algorítmicas opacas ao controle democrático.

Essa aproximação entre *Big Techs* e estruturas estatais revela um problema estrutural: a ausência de uma regulação clara e eficaz sobre os limites legais e éticos do uso de IA na esfera pública. Como argumenta Morozov, toda tecnologia incorpora uma racionalidade política: ela é desenhada, aplicada e operada segundo interesses específicos[25]. A inexistência de leis que disciplinam o uso de algoritmos em campanhas, a opacidade dos sistemas de recomendação e a falta de responsabilização por manipulação comportamental criam um vazio regulatório perigoso, no qual a liberdade política e a soberania cognitiva tornam-se vulneráveis à captura invisível.

Esse vazio regulatório impacta diretamente as garantias fundamentais para a proteção da mente humana frente às novas tecnologias de leitura e influência neural, entendidos como neurodireitos. A liberdade cognitiva, nesse cenário, é dissolvida não por coerção direta, mas por estratégias sofisticadas de condicionamento preditivo e modulação emocional automatizada. A democracia se converte, assim, em uma *performance* simbólica: as decisões parecem livres, mas foram previamente induzidas por estruturas algorítmicas que operam fora do alcance consciente do indivíduo.

Portanto, a eleição de Trump, em 2024, marca a consolidação de um modelo político baseado na manipulação cognitiva em larga escala. A IA não apenas interfere no espaço público: ela o redesenha, o condiciona e o gerencia segundo lógicas privatizadas. A liberdade política torna-se uma simulação, confinada nas bolhas informacionais desenhadas por sistemas que privilegiam engajamento em detrimento da verdade. Como adverte Zuboff, "a democracia não sobreviverá ao capitalismo de vigilância se não for capaz de enfrentá-lo com coragem e lucidez"[26].

---

[25] Evgeny Morozov, *To Save Everything, Click Here: The Folly of Technological Solutionism*, New York: PublicAffairs, 2013, p. 11.

[26] Shoshana Zuboff, *A era do capitalismo de vigilância…*, *op. cit.*, p. 514.

4. Entre o domínio e a emancipação: conceitos-chave em Hannah Arendt

O pensamento político clássico enfrenta um desafio complexo ao tentar compreender as transformações políticas e tecnológicas do século XXI. Ainda assim, os escritos de Hannah Arendt permanecem notavelmente atuais e pertinentes para a análise da realidade política contemporânea. Seu arcabouço teórico-conceitual oferece subsídios valiosos para entender as novas formas de dominação e a progressiva erosão da liberdade no contexto do capitalismo de vigilância. Nesse sentido, este capítulo propõe apresentar, de forma sintética, os principais conceitos desenvolvidos pela autora, com ênfase no totalitarismo, na banalidade do mal, na verdade factual e na liberdade enquanto ação política. Embora esses conceitos tenham sido elaborados no século XX, eles demonstram elevado potencial analítico para a compreensão das dinâmicas políticas atuais.

4.1. *O totalitarismo como forma*

Em *Origens do Totalitarismo* (1989), Arendt afirma que os regimes totalitários não devem ser confundidos com meras ditaduras convencionais, pois representam uma nova forma de governo, baseada em modos inéditos de dominação e opressão política e social[27].

Dessa forma, diferentemente das tiranias tradicionais, os regimes totalitários não se limitam ao domínio político da esfera pública, uma vez que almejam o controle total do pensamento e da consciência dos indivíduos. Tal dominação é alcançada por meio da manipulação sistemática da realidade, com o objetivo de disseminar o ideal totalitário até que este se torne onipresente. Nesse processo, os indivíduos passam a ser meros portadores da "lei do movimento", isto é, agentes inconscientes de uma lógica que opera independentemente da racionalidade ou da vontade humanas[28].

O totalitarismo, portanto, implica não apenas a destruição da esfera pública e do espaço de deliberação política, mas também a substituição da ação espontânea por comportamentos rigidamente condicionados[29]. Os sujeitos são, assim, reduzidos à mera funcionalidade dentro de uma engrenagem ideológica que anula sua capacidade de julgamento e iniciativa. Desse modo, a ideologia totalitária transforma o terror em mais do que um instrumento de repressão: ele

---

[27] Hannah Arendt, *Origens do totalitarismo*, Tradução de Roberto Raposo, São Paulo: Companhia das Letras, 1989, p. 7.

[28] *Ibidem*, p. 389.

[29] *Ibidem*, p. 322.

se converte na própria essência do regime, sustentado pelo controle psicológico e pela manipulação da realidade como mecanismos centrais de dominação[30].

4.2. *A banalidade do mal*

Em *Eichmann em Jerusalém: um relato sobre a banalidade do mal* (1999), Hannah Arendt formula o conceito de "banalidade do mal" a partir da análise do julgamento de Adolf Eichmann, oficial nazista responsável pela logística que viabilizou o extermínio de milhões de judeus durante o Holocausto. Nessa perspectiva, observa-se uma ruptura paradigmática em relação à concepção tradicional do mal como sinônimo de crueldade e perversidade extremas. Durante o julgamento, Arendt não identificou em Eichmann traços de monstruosidade moral, mas, sim, uma figura ordinária, um burocrata incapaz de refletir criticamente sobre seus atos, os quais executava de maneira mecânica e impessoal. A ausência de pensamento e de juízo moral permitiu que ele cometesse uma série de atrocidades sem considerar as consequências de suas ações[31].

O conceito de "banalidade do mal", portanto, não se refere a uma forma superficial ou atenuada de maldade, mas à possibilidade de que atos profundamente cruéis sejam praticados por indivíduos comuns, que não necessariamente agem por perversidade, mas por obediência cega a normas e ordens, sem questionar o conteúdo ético de suas ações. Assim, há uma desconstrução da visão tradicional do mal, uma vez que, nesse caso, a maldade apresentada por Eichmann revela-se "terrivelmente normal"[32]. Essa normalidade é fruto da renúncia à capacidade de julgar, da abdicação da racionalidade e dos valores morais em favor da conformidade com regras e hierarquias institucionais.

Importa frisar que Arendt não isenta Eichmann de responsabilidade; ao contrário, sua análise propõe uma nova compreensão do mal na modernidade. A "banalidade do mal" reside na incapacidade de pensar no sentido arendtiano, ou seja, de interromper a cadeia automática da obediência para refletir criticamente sobre o significado e os efeitos das próprias ações. Nesse sentido, o mal torna-se uma possibilidade concreta na vida moderna, sobretudo em contextos onde a burocracia e a tecnocracia substituem o julgamento ético individual e dissolvem a responsabilidade pessoal.

---

[30] *Ibidem*, p. 427.

[31] Hannah Arendt, *Eichmann em Jerusalém: um relato sobre a banalidade do mal*, Tradução de José Rubens Siqueira, São Paulo: Companhia das Letras, 1999, p. 165.

[32] *Ibidem*, p. 288.

4.3. *Verdade e a política*

No ensaio *Verdade e Política* (1967), Hannah Arendt analisa a tensão entre o domínio da verdade e o domínio da política, destacando-os como esferas regidas por lógicas distintas. No campo político, a verdade nem sempre é bem-vinda, uma vez que esse espaço opera por meio da persuasão, da opinião e da construção de narrativas que moldam o pensamento coletivo[33].

Para compreender plenamente essa reflexão, é fundamental distinguir os dois tipos de verdade abordados por Arendt: as verdades racionais – como aquelas produzidas pelas ciências – e as verdades factuais, que dizem respeito a acontecimentos concretos e verificáveis no mundo social. Estas últimas são especialmente vulneráveis à manipulação política, pois a distorção dos fatos colabora para minar os alicerces da realidade compartilhada[34].

Nesse sentido, o poder político frequentemente tenta apagar ou reconfigurar verdades factuais que possam representar uma ameaça à manutenção de determinadas narrativas hegemônicas. Esse processo de manipulação contínua acaba por substituir deliberadamente a realidade por ficções convenientes, gerando uma confusão generalizada entre verdade e engano. Com o tempo, instala-se uma desconfiança difusa: os indivíduos já não sabem em que acreditar e passam a duvidar de qualquer informação[35].

Consequentemente, essa crise da confiança compromete a própria possibilidade do espaço público. Para Arendt, a verdade factual é uma condição ontológica da política: sem um mundo comum de referências compartilhadas, o debate democrático se torna inviável. Afinal, a ação política só pode se desenvolver onde há uma realidade comum reconhecida por todos.

4.4. *Liberdade e ação pública*

Diferentemente das concepções clássicas de liberdade, centradas na ideia de livre-arbítrio e ausência de coerção, a concepção arendtiana propõe uma nova forma de compreender esse conceito. Em *A condição humana* (2007), Arendt apresenta a liberdade como ação, ou seja, como a capacidade de iniciar algo novo no mundo por meio da atuação no espaço público. Nesse sentido, a liberdade é

---

[33] Hannah Arendt, *Entre o passado e o futuro*, Tradução de Mauro W. Barbosa de Almeida, São Paulo: Companhia das Letras, 2009, p. 8.

[34] *Ibidem*, p. 15.

[35] *Ibidem*, p. 16.

indissociável da esfera pública, uma vez que só se realiza plenamente na convivência com outros indivíduos, na construção conjunta de um mundo comum[36].

Além disso, é importante destacar que a liberdade não pode ser reduzida a uma condição interior da vontade individual. Para Arendt, ela se manifesta concretamente apenas no espaço público, especialmente na imprevisibilidade da ação entre os homens. Nessa perspectiva, a liberdade, longe de ser soberania ou domínio de si, é a capacidade de iniciar algo novo e portanto, de transformar o mundo.

Assim, quaisquer mecanismos que inibam a espontaneidade da ação, dificultando a pluralidade e a deliberação pública, representam uma ameaça direta à liberdade. A imposição de estruturas e plataformas que promovam a uniformidade e a automatização da vida política contribui para a destruição da liberdade humana autêntica[37], uma vez que esta constitui a própria essência da política. Para que haja política, é necessário que exista espaço para a espontaneidade, a diversidade de discursos e a ação conjunta entre diferentes sujeitos, em prol da transformação do mundo.

5. Totalitarismo revisitado: perspectivas arendtianas para a era digital

A arquitetura algorítmica da inteligência artificial, ao transformar dados humanos em padrões previsíveis e manipuláveis, impõe uma nova forma de encapsulamento subjetivo que compromete diretamente a noção arendtiana de liberdade como ação pública. Em Hannah Arendt, a liberdade não é uma qualidade interior ou um direito abstrato, mas algo que emerge entre os homens, na pluralidade e na capacidade de iniciar algo novo no espaço compartilhado da política. No entanto, o ambiente digital contemporâneo, hoje amplamente dominado por sistemas de IA, desenhados e controlados por grandes corporações tecnológicas, sufoca essa pluralidade ao personalizar radicalmente a experiência de mundo. Cada indivíduo passa a habitar uma bolha informacional construída por critérios comerciais e algoritmos opacos, onde o que se vê, lê e pensa é continuamente filtrado e moldado por interesses que escapam ao debate público[38].

Ao contrário da promessa inicial da internet e das tecnologias digitais como espaços de livre circulação de ideias, a incorporação massiva de IA aos fluxos comunicacionais gerou ambientes de reforço cognitivo, isolando os sujei-

---

[36] Hannah Arendt, *A condição humana*, Tradução de Roberto Raposo, 10.ª ed., Rio de Janeiro: Forense Universitária, 2007, p. 189.

[37] *Ibidem*, pp. 248-249.

[38] *Ibidem*.

tos em nichos cada vez mais homogêneos e polarizados. A IA, ao predizer preferências e recomendar conteúdos, intensifica a criação de microcomunidades autorreferentes, nas quais as opiniões, em vez de dialogarem com visões diversas, ressoam apenas entre iguais, promovendo uma radicalização silenciosa e progressiva das posições. Assim, a IA não apenas reproduz as bolhas informacionais, mas as potencializa, eliminando a exposição ao conflito e ao debate público, elementos essenciais à democracia e à ação política, conforme Arendt[39].

Esse fenômeno representa não apenas uma distorção do espaço público, mas um risco concreto à liberdade política e ao pensamento reflexivo. A IA, enquanto promessa de automação e emancipação comunicacional, transformou-se em instrumento de confinamento subjetivo, conduzindo a uma nova forma de dominação que, embora sutil, compromete os fundamentos do agir político e da convivência democrática. As decisões automatizadas, a modulação comportamental e a hiperpersonalização da experiência cotidiana em plataformas digitais não só limitam a autonomia dos indivíduos como também obscurecem os critérios éticos e políticos sob os quais essas decisões são tomadas.

Nesse novo paradigma sociotécnico, é possível identificar uma expressão contemporânea daquilo que Hannah Arendt denominou "banalidade do mal". No julgamento de Adolf Eichmann, Arendt constatou que o executor logístico do Holocausto não era um fanático ideológico ou psicopata, mas um funcionário comum, obediente, incapaz de reflexão crítica sobre suas ações. O mal manifestava-se não por intenções perversas, mas pela ausência de pensamento, sustentada pela repetição mecânica de ordens e procedimentos administrativos[40].

Analogamente, na era da inteligência artificial, assistimos à emergência de uma banalização automatizada do mal: uma engrenagem impessoal composta por algoritmos, códigos e plataformas que moldam percepções, induzem comportamentos e reconfiguram o espaço público sem mediação crítica ou responsabilidade visível. A dinâmica algorítmica privilegia conteúdos que geram engajamento, sensacionalistas, polarizadores ou fraudulentos, não por escolha deliberada de agentes malignos, mas porque tais conteúdos são eficazes para manter o usuário conectado e gerar lucro para as corporações controladoras.

No modelo atual, todos são simultaneamente emissores e receptores de conteúdo, inseridos em uma lógica de retroalimentação algorítmica na qual as máquinas se nutrem das reações humanas para calibrar suas próximas sugestões. Conteúdos agressivos, odientos e mentirosos tendem a ser amplificados não por uma decisão consciente, mas porque são os mais eficazes para maximizar a permanência do usuário nas plataformas. Assim, ao interagir com esses conteúdos,

---

[39] Hannah Arendt, *Entre o passado e o futuro*, *op. cit.*, p. 8.

[40] Hannah Arendt, *Eichmann em Jerusalém: um relato sobre a banalidade do mal*, *op. cit.*, p. 165.

mesmo que passivamente, os próprios usuários tornam-se vetores involuntários da sua difusão, participando de um ecossistema que banaliza o mal e dissolve os critérios éticos do discurso público. É como se estivéssemos diante de milhões de "Eichmanns digitais": não indivíduos conscientes de sua maldade, mas agentes funcionais de um sistema que opera sem pensamento crítico[41].

A banalidade do mal digital e algorítmico não exige convicções ideológicas ou ódio deliberado; basta a repetição acrítica de comportamentos e a aceitação naturalizada de um ambiente projetado para operar sem consciência ética. O dano se perpetua de forma silenciosa, incorporado nos dados coletados, nas recomendações personalizadas, nos vieses confirmados e reforçados por sistemas supostamente neutros, mas moldados por interesses comerciais opacos.

O século XXI testemunha, portanto, a ascensão de uma nova forma de mal: técnica, funcional, silenciosa, mas profundamente corrosiva. Diferente dos regimes totalitários tradicionais, que exigiam a imposição autoritária da ideologia, agora o mal se insinua sob a promessa de liberdade, conexão e personalização, dissolvendo a ação pública reflexiva e enfraquecendo a pluralidade sobre a qual se funda a vida democrática.

Esse cenario teve seu ápice paradigmático durante a campanha eleitoral de Donald Trump, em 2024, na qual, semelhantemente às eleições de 2016, a inteligência artificial foi utilizada para manipular emocionalmente os indivíduos. A disseminação de conteúdos sintéticos, *deepfakes* e notícias que não representavam a realidade factual teve o intuito claro de dominar as massas por meio do medo e da confusão, propagados pelas arquiteturas algorítmicas das plataformas digitais[42]. Sob essa ótica, a eleição de Trump, em 2024, não pode ser vista como mero fenômeno de *fake news*, mas como a institucionalização de um cenário de desconfiança total, que corrobora para a degradação do ambiente democrático.

A crença nessas narrativas falaciosas decorre de um estado de passividade política induzida pela manipulação algorítmica, pois a exposição repetitiva a notícias distorcidas supera a capacidade de julgamento autônomo dos indivíduos. Quando a verdade factual é negada sistematicamente, não se cria uma nova crença, mas se corrói o próprio senso de realidade da sociedade[43]. Assim, a eleição de Donald Trump confirma a lógica de que a verdade factual não importa; o que importa é a mobilização e o engajamento dos indivíduos, inaugurando um governo que não atua sobre o real, mas sobre o simbólico.

A consolidação dessa nova racionalidade de poder não depende mais de um acontecimento específico para se legitimar, pois opera de maneira difusa,

---

41 *Ibidem*.

42 Hannah Arendt, *Entre o passado e o futuro*, *op. cit.*, p. 275.

43 *Ibidem*, p. 278.

contínua e autorreferente. O que se observa não é apenas a instrumentalização política do ambiente informacional, mas a internalização de sua lógica: a política passa a existir menos como espaço de confronto público e mais como arena de gestão de percepções, em que narrativas não disputam a verdade, mas a intensidade de adesão emocional que são capazes de produzir. Nesse modelo, a autoridade política não deriva da argumentação ou do consenso social, mas da capacidade de coordenar fluxos de atenção, modular sentimentos coletivos e converter engajamento em suporte político estável, independentemente da correspondência com a realidade empírica.

Esse deslocamento evidencia que o ponto central da transformação em curso não é um líder ou evento isolado, mas a própria infraestrutura de mediação social, que reorganiza as relações entre governantes, governados e opinião pública. Quando a política se torna dependente dos mecanismos de circulação e retroalimentação algorítmica, a esfera pública perde sua função deliberativa e assume um caráter performático: ela não produz cidadãos, produz audiências; não articula ação coletiva, articula impulsos coletivos; não organiza consensos negociados, mas alinhamentos comportamentais mensuráveis. Assim, a ameaça à democracia não se expressa apenas no conteúdo que circula, mas na gramática que o torna possível, uma gramática que dissolve a distinção entre participação política e condicionamento psicológico sutil, corroendo a própria possibilidade de um espaço público autônomo.

Como desdobramento dessa transformação da política em um ambiente de gestão algorítmica da atenção e modulação do comportamento, o espaço público passa a ser ocupado por arquiteturas privadas de poder informacional, que não se organizam a partir de princípios de publicidade, controle social ou responsabilidade democrática. Isso representa a ocupação dos espaços públicos por estruturas privadas que atuam sem transparência e sem compromisso ético, convertendo a política em um espetáculo dirigido[44].

O ideal democrático clássico torna-se impraticável diante de uma estrutura sistêmica de dominação algorítmica que inviabiliza a diferença de opiniões, reduz o debate a simplismos e condiciona a participação política à lógica do algoritmo. Assim, a democracia é substituída por uma simulação participativa, na qual os cidadãos são meros usuários das plataformas digitais e a liberdade de ação é convertida em comportamento rastreável pelas *Big Techs*.

Nesse cenário, o governo Trump representa uma nova forma de exercício do poder, na qual a mentira não é apenas aceita, mas institucionalizada; a ação política é substituída por *performances* simbólicas; e a verdade factual é moldada pelos interesses das plataformas digitais e algoritmos alheios à ética pública e ao

[44] Hannah Arendt, *A condição humana*, *op. cit.*, p. 244.

ideal democrático. Esse contexto intensifica a construção de regimes totalitários reconfigurados, nos moldes arendtianos.

A realidade política contemporânea evidencia, assim, uma reconfiguração do totalitarismo moderno, em que a esfera pública é substituída por uma lógica de gestão algorítmica operada pelas *Big Techs*. Esse processo constitui um colonialismo psíquico digital, que não depende mais das instituições repressoras clássicas para exercer dominação, mas atua diretamente sobre a capacidade de julgar, pensar e agir dos indivíduos. Trata-se de uma colonização da interioridade humana, realizada pela coleta massiva de dados pessoais e pela modulação algorítmica hiperpersonalizada da experiência, comprometendo profundamente a liberdade individual.

Essa colonização manifesta-se na simbiose entre o governo Trump e as grandes plataformas digitais, que se tornaram instrumentos de propaganda e gestão da opinião pública, incentivando a despolitização do espaço democrático e a homogeneização das condutas individuais. Sob essa perspectiva, o totalitarismo digital não opera pela repressão explícita, mas pela normatização algorítmica e pela fabricação de consensos que promovem uma aparência de liberdade e democracia, enquanto bloqueiam as condições essenciais para o exercício pleno da cidadania.

O aniquilamento do espaço político, um dos pilares centrais dos regimes totalitários descritos por Arendt, ocorre hoje não por censura deliberada, mas pela fragmentação informacional em grupos inconciliáveis. A política se torna arena de crenças individuais e fictícias, resultando em isolamento cognitivo, condição essencial para o domínio totalitário[45]. Indivíduos em bolhas informacionais, privados de interlocução plural, tornam-se vulneráveis à dominação de estruturas privadas com controle do poder informacional.

Dessa forma, a configuração contemporânea do poder revela uma mutação paradigmática do totalitarismo moderno: já não é um regime opressor clássico, mas uma dominação silenciosa e psicológica. Esse totalitarismo reconfigurado controla a totalidade da vida humana por meio da ausência deliberada de regulação das *Big Techs*, que seguem modulando preferências e condicionando preventivamente a ação política. Trata-se de um projeto que não apenas controla o Estado, mas molda a consciência coletiva e corrói estruturalmente a liberdade individual.

Diante desse cenário de manipulação algorítmica, desintegração da verdade factual e institucionalização da mentira, o Direito é convocado com urgência. Não basta regular plataformas em aspectos comerciais ou concorrenciais; é necessário enfrentá-las como estruturas de poder que afetam diretamente a

[45] Hannah Arendt, *Origens do totalitarismo*, *op. cit.*, p. 701.

liberdade mental e a autodeterminação dos indivíduos. A arquitetura algorítmica não é neutra; ela molda o que se vê, sente e pensa. Por isso, não basta falar em privacidade ou responsabilidade civil: é urgente inaugurar uma nova geração de direitos fundamentais, os neurodireitos, focados na proteção da mente humana.

Portanto, proteger a integridade cognitiva, garantir a soberania da consciência e resguardar o pensamento autônomo são tarefas centrais de um novo constitucionalismo ético e tecnológico. Em um tempo em que o poder volta-se diretamente à mente, o silêncio do Direito equivale à conivência. Regular a inteligência artificial não é limitar a liberdade, mas sua condição de possibilidade no século XXI. Nesse horizonte, desenha-se a urgência de uma agenda normativa que enfrente, com coragem e lucidez, os riscos do totalitarismo reconfigurado.

Conclusão

Na interseção entre controle e consciência, os neurodireitos surgem como uma resposta jurídica e ética à invasão silenciosa da mente humana realizada por sistemas de inteligência artificial, plataformas digitais e arquiteturas algorítmicas que moldam emoções, percepções e decisões. O que está em jogo não é apenas a privacidade dos dados ou a transparência dos algoritmos, mas a própria essência da liberdade de pensar, escolher e existir. À medida que as tecnologias digitais avançam além da comunicação e começam a interferir diretamente nos processos mentais e neuropsicológicos que fundamentam nossas decisões, torna-se claro que os marcos tradicionais dos direitos humanos são insuficientes diante das novas vulnerabilidades cognitivas.

A ascensão de regimes tecnopopulistas, como o segundo mandato de Donald Trump, sustentada por estratégias sofisticadas de manipulação comportamental e engajamento emocional algorítmico, evidencia os perigos de uma democracia reduzida a uma simulação participativa, onde o cidadão se torna mero usuário e as decisões políticas são condicionadas por estímulos digitais meticulosamente projetados. O século XXI exige não apenas a proteção do corpo físico contra a violência estatal, mas também a defesa do espaço mental contra a manipulação invisível e constante de inteligências artificiais que moldam preferências, influenciam escolhas e distorcem o julgamento individual.

Nesse novo contexto, os neurodireitos não podem ser vistos como uma especulação futurista ou um debate marginal. Devem ser encarados como um imperativo civilizacional diante de um presente marcado pela erosão da autonomia e pelo risco de sermos transformados em engrenagens de um sistema que opera por meio da vigilância, do condicionamento e da captura emocional. Proteger a mente humana contra a manipulação tecnológica é, portanto, proteger o núcleo mais íntimo da dignidade, da liberdade e da democracia.

A engenharia algorítmica não atua apenas sobre os dados que produzimos, mas sobre nossa capacidade de julgamento. Quando as plataformas moldam a forma como percebemos o mundo, interferem diretamente na formação da consciência crítica. O resultado é um ambiente onde a razão dá lugar à emoção, o discernimento cede espaço à impulsividade, e a política se transforma num espetáculo dirigido por interesses privados. Desinformação, polarização e radicalismo não são falhas aleatórias do sistema, mas sintomas de uma arquitetura projetada para capturar atenção, modular comportamento e gerar lucro, mesmo que isso comprometa o próprio tecido democrático.

Diante desse cenário, assegurar a autodeterminação cognitiva torna-se a tarefa central do Direito na era da inteligência artificial. Isso significa reconhecer que a liberdade mental é a condição essencial para todas as demais liberdades. Significa afirmar que a mente humana não pode ser tratada como objeto de exploração técnica nem alvo de experimentação comportamental. Sobretudo, significa entender que a regulação das tecnologias emergentes não limita a liberdade, mas a torna possível.

Frente aos riscos concretos de manipulação sistêmica, desinformação orquestrada e redução da subjetividade à previsibilidade algorítmica, os neurodireitos se configuram como a fronteira normativa capaz de proteger o humano dentro do digital. Garantir a integridade cognitiva, preservar a soberania da consciência e reafirmar o direito de cada indivíduo à autoria de si mesmo não é apenas um desafio jurídico: é uma missão histórica. Trata-se, acima de tudo, da luta para que a humanidade não se perca diante do fascínio sedutor da inteligência artificial.

SER IDOSO NO SÉCULO XXI: DESAFIOS E PERSPETIVAS SOBRE O DIREITO À PROTEÇÃO DA SAÚDE NA ERA DIGITAL

*Sofia Ferreira da Costa**

Resumo: O presente artigo visa a análise e o impacto dos sistemas de Inteligência Artificial no acesso à saúde da população idosa. O tema a tratar é particularmente relevante, considerando as últimas décadas de desenvolvimento tecnológico. Aliado a este facto, surgimos enquanto país como um dos mais envelhecidos do mundo e com um índice de envelhecimento que tem vindo a aumentar ao longo dos últimos anos. Em 2024, a esperança média de vida à nascença em Portugal apontava para os 84 anos para o sexo feminino e 78,7 para o sexo masculino, sendo de notar o sucessivo aumento ao longo das últimas décadas. Surgem, então, diversos desafios em diferentes áreas, nomeadamente quanto à crescente digitalização dos recursos e aos diminutos níveis de literacia digital na referida faixa etária que, não raramente, constituem verdadeiras adversidades.

* Doutoranda em Ciências Jurídicas Privatísticas na Escola de Direito da Universidade do Minho, Investigadora do JusGov – Centro de Investigação em Justiça e Governação e Advogada. E-mail: sofiaferreiradacosta-71089p@adv.oa.pt.

A *Inteligência* Artificial pode ser um forte aliado na melhoria dos cuidados de saúde, já que é capaz de processar uma grande quantidade de dados e fazer previsões baseadas nos padrões encontrados, sendo que as soluções são cada vez mais inovadoras à medida que a tecnologia também avança.

Não obstante as vantagens inerentes a este fenómeno, surgem dificuldades associadas, como questões relacionadas com o cariz discriminatório dos algoritmos, segurança e proteção de dados – além das referidas dificuldades de adaptação por parte da população idosa. Ocupar-nos-emos ainda das questões associadas à responsabilidade de decisões tomadas por entes dotados de Inteligência Artificial.

Palavras-chave: Idosos; Inteligência Artificial; Etarismo.

Abstract: This article aims to examine the impact of Artificial Intelligence systems on the access to healthcare by the elderly population. The topic is particularly relevant in light of the technological advancements of the recent decades. Concurrently, Portugal has emerged as one of the most aged countries in the world, with an ageing index that has steadily increased over the recent years. In 2024, life expectancy at birth in Portugal was estimated at 84 years for women and 78.7 years for men, reflecting a continuous upward trend.

This demographic evolution presents significant challenges across multiple domains, particularly regarding the increasing digitalisation of services and the low levels of digital literacy within this age group, which often represent a considerable barrier to access.

Artificial Intelligence holds considerable potential to enhance healthcare provision, given its capacity to process large volumes of data and to generate predictions based on identified patterns. These solutions are becoming progressively more innovative as the technology evolves.

Notwithstanding, alongside these benefits, several challenges persist. These include concerns related to algorithmic bias, data security and privacy, as well as the aforementioned difficulties in adaptation among the elderly. This article also addresses the critical issue of accountability for the use of AI and for decisions made by autonomous systems.

Keywords: Elderly; Artificial Intelligence; Ageism.

1. A pessoa idosa e o direito à proteção na saúde

O envelhecimento populacional é um fenómeno atual e, assim sendo, será necessário adotar medidas em conformidade com este circunstancialismo. A Constituição da República Portuguesa (doravante, CRP) dedica o seu artigo 72.º à "terceira idade", manifestando preocupações com o seu bem-estar e, designadamente, com a sua marginalização social, advogando por "oportunidades de realização pessoal, através de uma participação activa na vida da comunidade". Também no artigo 64.º, sob a epígrafe "Saúde", surge o direito à proteção da mesma, com o consequente dever de a defender e promover. Note-se, por um lado, o direito de exigir ações e prestações por parte dos serviços públicos e, por outro lado, o dever geral de abstenção de práticas que degradem ou afetem negativamente a saúde humana[1]. Note-se ainda o disposto no artigo 67.º da Lei Fundamental que confere proteção à família enquanto elemento fundamental da sociedade. No mesmo sentido aponta a Carta de Direitos Fundamentais da União Europeia que "reconhece e respeita o direito das pessoas idosas a uma existência condigna e independente e à sua participação na vida social e cultural", nos termos do artigo 25.º. Note-se ainda o vertido no artigo 35.º do suprarreferido diploma, que prevê a faculdade a "todas as pessoas" de "aceder à prevenção em matéria de saúde e beneficiar de cuidados médicos, de acordo com as legislações e práticas nacionais".

É de elementar importância, considerando o aumento sucessivo da esperança média de vida, estudar e analisar as condições existentes e as eventuais necessidades que surjam com os desafios suscitados[2]. Foi neste sentido que a Assembleia Geral das Nações Unidas estabeleceu princípios para as pessoas idosas, adotados pela Resolução 46/91, de 16 de dezembro de 1991. Instituem-se, por isso, cinco princípios basilares, sendo eles a independência, participação, assistência, realização pessoal e dignidade. Trata-se de uma importante demonstração da necessidade de considerar as definições políticas na saúde, sobretudo sendo o envelhecimento populacional um fenómeno que tem afetado particularmente a Europa, desde os finais do século passado[3]. Se consideramos e atendemos à necessidade de proteção de crianças e jovens, a mesma vulnerabilidade poderá também ser encontrada nos idosos, pelo que, para esta faixa etária, igualmente existem situações de perigo. Têm sido, assim, debatidos na doutrina

---

[1] Maria João Estorninho, Tiago Macieirinha, *Direito da Saúde*, Universidade Católica, Lisboa, 2014, p. 45.

[2] No início do século passado, a esperança média de vida à nascença rondava os 47 anos, quer no Reino Unido, quer nos Estados Unidos da América. Nos últimos anos – em pouco mais de 100 anos – esse registo já ultrapassou os 80 anos. Veja-se https://rpmgf.pt/ojs/index.php/rpmgf/article/view/13515 [consultado em 16/09/2025].

[3] Cláudia Monge, «A proteção e a promoção da saúde da pessoa idosa», in Carla Amado Gomes, Ana F. Neves (coords.), *Direito e direitos dos Idosos*, AAFDL, Lisboa, 2020, pp. 133-180.

a importância e o papel do Direito da Família na tutela da pessoa idosa[4]. Note-se que será um conceito cuja determinação não é unânime ou consensual. À luz do direito português, poderemos exemplificar com o disposto na alínea b) do n.º 1 do artigo 1720.º do Código Civil, que estabelece a imperatividade do regime de bens para casamentos celebrados por quem tenha mais de 60 anos. Por outro lado, na vida prática, sabemos que é a partir dos 65 anos que surgem distinções, nomeadamente no preço dos transportes públicos ou no benefício de políticas como o apoio social para idosos[5].

O envelhecimento e a qualidade de vida associada ao mesmo serão por isso fatores a considerar no âmbito do Direito da Saúde, sendo que, para tal, consideraremos que este ramo integra "todo o complexo normativo cujo objeto principal é a proteção e a promoção da saúde"[6]. Para este contexto importa estabelecer a distinção entre o que se entende por obrigações de meios e de resultado, visando resolver o problema associado ao ónus de prova da responsabilidade civil médica que abordaremos *infra*.

2. A digitalização dos recursos nos cuidados de saúde

No que à digitalização dos cuidados de saúde concerne, Portugal tem vindo a destacar-se como sendo um dos países mais avançados nesta área ao nível da União Europeia (doravante, UE). O Relatório designado como "[D]igitalização dos cuidados de saúde" visa analisar o progresso existente nos Estados-Membros, tendo Portugal alcançado o décimo lugar quanto ao acesso a dados de saúde. Quanto à maturidade dos serviços em linha relacionados com saúde, Portugal alcançou 79 pontos percentuais, o que garantiu o acesso ao sexto lugar da lista encabeçada pelo Luxemburgo e Malta, com um empate de 95%. Portugal encontra-se, assim, acima da média da UE quando são comparados os dados da Administração Pública em linha e o indicador de saúde em linha da década digital, tendo estes dados sido melhorados ao longo dos últimos anos, só podendo o cenário ser, por isso, favorável[7].

---

[4] Vítor Palmela Fidalgo, «A tutela do idoso no Direito da Família», in Carla Amado Gomes, Ana F. Neves (coords.), *Direito e direitos dos Idosos*, AAFDL, Lisboa, 2020, p. 334.

[5] Decorrem debates sobre a idade em que alguém deve ser considerado "idoso", sobretudo nos países mais afetados pelo envelhecimento. No Japão, as Sociedades Gerontológicas e Geriátricas questionam o limiar dos 65 anos, apresentando argumentos que redefinem o conceito em questão. Veja-se https://cenie.eu/pt/noticias/velhice-ja-nao-comeca-aos-65-anos [consultado em 16/09/2025].

[6] Maria João Estorninho, Tiago Macieirinha, *Direito da Saúde*, *op. cit.*, p. 16.

[7] Disponível em https://www.eca.europa.eu/ECAPublications/SR-2024-25/SR-2024-25_PT.pdf [consultado em 22/09/2025].

Atualmente, podemos considerar a "Saúde Digital" ou "eSaúde" – (do inglês *eHealth*) – como um fenómeno tipicamente associado ao presente século, podendo ser entendido como "um conjunto de ferramentas e serviços que utilizam as tecnologias de informação e comunicação 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"[8]. A verdade é que a utilização da tecnologia nesta área tem sido cada vez mais recorrente, desde funcionalidades mais básicas, como o agendamento *online* de uma consulta – ou até mesmo a uma teleconsulta –, até à introdução dos mais modernos e avançados sistemas de Inteligência Artificial (doravante, IA) que permitem, por exemplo, realizar cirurgias e até mesmo cirurgias à distância através da utilização de rede 5G e recurso à IA – telecirurgia[9]. Este fenómeno poderá ser, no futuro, essencial para situações de catástrofe natural ou de socorro em cenários bélicos. As vantagens associadas parecem evidentes, não obstante, surgem também aspetos que devem justificar uma análise profunda e ponderada, nomeadamente, os recursos humanos necessários para tal, bem como recursos económicos investidos nos materiais e equipamentos e o impacto ambiental em causa[10].

Ao longo dos anos, a tecnologia tem avançado significativamente, o que naturalmente proporciona o surgimento de novas soluções que podem ser implementadas. No âmbito da saúde, têm-se desenvolvido mais e mais precisas técnicas e métodos de diagnóstico, também impulsionados pelo constante armazenamento de dados e pela aprendizagem autónoma com base em experiên-

---

[8] Disponível em https://health.ec.europa.eu/ehealth-digital-health-and-care/digital-health-and-care_pt [consultado em 22/09/2025].

[9] A primeira cirurgia realizada com a intervenção de um robot – PUMA 560 – data de 1985, tendo sido utilizado para uma biópsia cerebral. Em 2024, foi realizada uma telecirurgia que impulsionou por completo esta inovação cirúrgica. Tratou-se de uma operação em que o médico se encontrava em Roma e o paciente na China. Veja-se https://www.britannica.com/technology/PUMA-560; https://www.scmp.com/news/china/science/article/3266069/chinese-doctor-performs-telesurgery-first-rome-beijing-procedure [consultado em 22/09/2025].

[10] O crescente uso de mecanismos de Inteligência Artificial pode ser verdadeiramente preocupante em vários aspetos, sendo a pegada ambiental um problema associado ao consumo de recursos naturais como água, energia e recursos minérios. Os *data centres* onde são armazenadas grandes quantidades de dados e as matérias minerais associadas às componentes eletrónicas são produtores de grandes quantidades de lixo eletrónico. Estima-se que, para construir um computador com cerca de 2kg, são necessários cerca de 800kg de matéria-prima. O calor produzido e a energia obtida a partir da queima de combustíveis fósseis proporciona também a libertação de gases que, por sua vez, aumentam o efeito de estufa. Como mecanismo de aquecimento destes sistemas são necessárias grandes qauntidades de água potável, fator particularmente alarmante quando cerca de um quarto do planeta não dispõe deste recurso. A Organização das Nações Unidas criou um programa para o meio ambiente (PNUMA) em que alerta para a problemática do impacto da IA no meio ambiente e estabelece algumas recomendações para minimizar o impacto da mesma. Veja-se https://www.unep.org/news-and-stories/story/ai-has-environmental-problem-heres-what-world-can-do-about [consultado em 22/09/2025].

cias anteriores que caracteriza o desenvolvimento dos sistemas de IA (*machine learning*). Este comportamento permite a análise de uma grande quantidade de dados e o desenvolvimento de padrões que permitem um diagnóstico mais correto e preciso[11]. Todas estas vicissitudes garantem uma melhor experiência e tratamento ao paciente[12], bem como aos profissionais de saúde – cujas decisões podem ser tomadas de forma mais rápida e eficiente. Sobre as consequências e responsabilidade destas decisões analisaremos *infra*.

2.1. *Da necessidade de uma literacia digital efetiva à exclusão digital dos idosos: um obstáculo à efetivação de direitos?*

O processo de adaptação ao novo contexto digital pode, efetivamente, ser um grande desafio para faixas etárias mais avançadas. Tratam-se de desafios em diversas vertentes, nomeadamente a nível cultural, social e de segurança[13]. Foram já aqui analisadas as diversas e inequívocas vantagens associadas aos sistemas de IA em várias vertentes, como diagnósticos mais precisos e mais rápidos, análise de exames de modo mais eficiente, entre outros. Não obstante, sabemos também – e será necessário estar consciente desse facto – do cariz discriminatório que caracteriza este tipo de sistemas. Surge, por isso, de forma inevitável o debate ético relativo ao uso de IA e algumas das questões que essa mesma IA coloca ao Direito. A capacidade de recolha e análise de dados e o seu tratamento constituem, de facto, um desafio para diversos Direitos Fundamentais e para os direitos relativos à tutela da personalidade, previstos no artigo 70.º do Código Civil.

Sucede que, na tomada de decisões, os algoritmos baseiam-se no padrão e na experiência existentes, muitas vezes reproduzindo resultados discriminatórios para com determinados grupos da sociedade, social ou economicamente menos favorecidos. Atualmente, as relações humanas encontram-se rodeadas de algoritmos, ainda que nem sempre sejamos capazes de nos apercebermos de tal facto. Com pesquisas nos motores de busca, é através dos algoritmos

---

[11] Veja-se https://kaizen.com/pt/insights-pt/transformacao-digital-cuidados-saude-inovacao/, com última visualização a 22/09/2025.

[12] De acordo com o "Eurobarómetro" lançado pela Comissão Europeia, 76% dos cidadãos da UE consideram que a conectividade avançada e a cibersegurança melhorarão a sua relação diária com as tecnologias digitais. Para este efeito, deve salientar-se que, entre os cidadãos da UE, apenas metade consideram que os seus direitos digitais estão bem protegidos. Veja-se https://europa.eu/eurobarometer/surveys/detail/2959?etrans=pt [consultado em 22/09/2025].

[13] No período de isolamento em contexto da pandemia COVID-19, o Japão desenvolveu um robot com o objetivo de fazer companhia, sobretudo a idosos em lares. Tratou-se de uma forma de combater o isolamento e a solidão. Veja-se https://pt.euronews.com/2021/01/04/o-japao-usa-a-tecnologia-para-atenuar-a-solidao-causada-pela-pandemia [consultado em 23/09/2025].

que podemos produzir resultados[14]. Tratam-se de procedimentos lógico-matemáticos que, com passos discretos, são eficazes na solução dos problemas a que se propõem[15]. Na sua dimensão técnica e computacional, podemos definir algoritmos como sendo "regras, formais, sequenciais, e bem definidas que a partir de um entendimento lógico-matemático se debruçam sobre um problema, possível de ser tratado e executado por um computador, em que dados de entrada são transformados em dados de saída"[16]. Podemos afirmar que os sistemas inteligentes visam replicar o pensamento humano de forma mais rápida e eficaz. Significará isto que os preconceitos e considerações humanas existentes serão inevitavelmente refletidos nas decisões tomadas, ainda que impliquem escolhas discriminatórias ou que ponham em causa a igual dignidade social que a todos é atribuída. Tal pode dever-se à ausência de condicionantes emocionais e falta de interesses do sistema computacional, não refletindo, por isso, decisões neutras[17]. Pode, por isso, surgir o problema associado ao facto de os algoritmos se apoiarem em determinados dados e poderem refletir, quer de forma premeditada ou não, preconceitos de diversos níveis[18]/[19].

Será, por isso, essencial lembrar que a idade não poderá ser motivo para colocar em questão os direitos de personalidade e Direitos Fundamentais dos cidadãos, até porque se tem vindo a demonstrar que o funcionamento algorítmico por via de juízos indutivos e dedutivos não transmite de modo seguro a qualidade da informação que é divulgada[20]. Os direitos de personalidade encontram-se tutelados nos artigos 70.º e ss. do Código Civil, sendo que, nas palavras de Carlos Alberto da Mota Pinto, tratam-se de "direitos gerais (todos deles gozam), extrapatrimoniais (embora as suas violações possam originar uma reparação em dinheiro, não têm, em si mesmo, valor pecuniário) e absolutos"[21]. Há

---

[14] Paulo Victor Alfeio Reis, *Algoritmos e o direito*, Almedina, São Paulo, 2020, p. 102.

[15] *Ibidem*, p. 107.

[16] *Ibidem*, p. 119.

[17] Mafalda Miranda Barbosa, *Inteligência Artificial – Entre a Utopia e a Distopia*, Gestlegal, Coimbra, 2024, p. 200.

[18] Miguel Ángel Presno Linera, *Derechos Fundamentales e Inteligencia Artificial*, Marcial Pons, Madrid, 2022, p. 26.

[19] No âmbito penal, concluiu-se que a utilização algorítmica pelas entidades policiais e judiciais afloram os preconceitos que podem estar inerentes a um conjunto de dados subjacentes utilizados para o efeito, nomeadamente dados históricos introduzidos e desenvolvidos por algoritmos. Estes preconceitos tendem, por isso, a exponenciar a dimensão discriminatória e, em particular, aqueles que que pertencem a grupos minoritários, étnicos ou radicalizados. Veja-se Miguel Ángel Presno Linera, *Derechos Fundamentales e Inteligencia Artificial*, *op. cit.*, pp. 26-27.

[20] Mafalda Miranda Barbosa, *Inteligência Artificial – Entre a Utopia e a Distopia*, *op. cit.*, p. 187.

[21] Carlos Alberto da Mota Pinto, *Teoria Geral do Direito Civil*, 4.ª ed., Coimbra Editora, Coimbra, 2012, pp. 208-209.

um círculo de direitos que podemos considerar para o presente caso, sendo eles a integridade moral e física, imagem, reserva sobre a intimidade da vida privada e a própria saúde física. A violação destes direitos, além de ilícito penal concretamente aplicável, pode ainda consistir num ilícito civil, como aliás estudaremos *infra*. A cláusula geral de personalidade plasmada no artigo 70.º do Código Civil não deverá ser atendida sem o regime constitucionalmente consagrado de Direitos, Liberdades e Garantias a que correspondem direitos de personalidade[22]. Os direitos de personalidade, pela essencialidade que revestem, são, por isso, inalienáveis e irrenunciáveis[23]. No que respeita à saúde, a falta de literacia digital nas gerações mais avançadas pode ser um fator – não exclusivo das gerações mais velhas, note-se – para que estes direitos possam ser postos em causa. Desde logo porque, estando impedido de conhecer os termos ou as condições associadas a um dado serviço, ficará naturalmente em causa o consentimento livre e esclarecido, podendo abrir-se caminho para o atropelo de uma série de outros direitos, nomeadamente os associados à recolha de dados pessoais, agravando-se pelo caráter sensível dos dados de saúde. Esta é uma preocupação expressa no Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho de 13 de junho de 2024 – *AI Act* –, que considera os sistemas de IA que apresentem riscos "para a saúde ou a segurança ou para os direitos fundamentais das pessoas" como sistemas que apresentam um risco. Estipula, por isso, o Regulamento que, nestas situações, o operador tome as medidas adequadas para assegurar que, quando colocado novamente no mercado, o sistema de IA já não mais apresenta esse risco. É ainda reconhecida a importância de dotar as pessoas para as noções necessárias para que, por esta via, possam tomar decisões informadas sobre os sistemas de IA. Com este investimento pretende-se proporcionar todos os conhecimentos para a correta utilização da IA, bem como contribuir para o aumento da confiança na mesma, compreendendo uma sensibilização para os riscos e benefícios à mesma associados, promovendo a adoção de códigos de conduta. Pretende-se que, concretamente no setor da saúde, os sistemas de diagnóstico e que apoiam decisões humanas estejam cada vez mais sofisticados, de forma a que surjam resultados exatos e, mais uma vez, de confiança. Ainda a propósito do cariz discriminatório dos algoritmos, refere-se que os resultados gerados pela IA poderão ser "influenciados por enviesamentos inerentes que tendem a aumentar gradualmente e, desse modo, a perpetuar e a ampliar a discriminação existente". Se as decisões são baseadas em decisões humanas anteriores, é elevada a probabilidade de esses mesmos erros se reproduzirem e multiplicarem, pelo que deveria ser a todos garantido o direito de um resultado de busca impar-

[22] *Ibidem*, p. 211.

[23] *Ibidem*, p. 215.

cial e não personalizado[24]. Assim, pretende-se evitar que decisões atentatórias da dignidade da pessoa humana baseadas na idade ou na esperança média de vida sejam adotadas, sendo também uma preocupação expressa no artigo 4.º do Regulamento a necessidade de implementação de medidas de investimento em literacia no domínio da IA face aos prestadores e responsáveis pela implantação de sistemas de IA.

2.2. *Discriminação e Etarismo*

Não obstante o princípio da igualdade e não discriminação constitucionalmente consagrado, tutelando a igualdade de todos os cidadãos perante a lei, a verdade é que, no domínio dos sistemas inteligentes, tal nem sempre sucede. Surgem também novos desafios associados ao aumento da esperança média de vida, sendo um fenómeno de realçar a nível mundial. Versando o presente artigo nos direitos dos idosos, é essencial abordar a questão do etarismo – *ageism*. Podemos definir o termo como a atribuição à idade de uma relevância tal que, com base nela, são tomadas decisões que podem ser desvantajosas, desfavoráveis ou injustas. O princípio da proteção da saúde, previsto no artigo 64.º da Constituição, estipula a proibição de discriminações ilegítimas de doentes, pelo que o controlo da igualdade tem uma importância elevada na fiscalização do modo como é garantido o acesso à proteção e aos cuidados de saúde[25]. A Organização Mundial de Saúde (doravante, OMS) entende que algumas escolhas da IA baseadas em algoritmos se baseiam em "estereótipos, preconceitos e discriminações ou a si mesmo baseado na idade"[26].

Estudos têm vindo a estabelecer a relação do aumento da idade com a propensão à discriminação[27]. Estima-se que as ocasiões mais propensas à ocorrência de tal fenómeno são em contexto laboral, serviços públicos, serviços de saúde, na família, nas instituições de apoio à terceira idade, entre outros[28]. Assim, forçoso será concluir que, para que os sistemas de IA alcancem os objetivos

---

[24] Miguel Ángel Presno Linera, *Derechos Fundamentales e Inteligencia Artificial*, *op. cit.*, p. 66.

[25] Rui Medeiros, «Artigo 64.º», in Jorge Miranda, Rui Medeiros (coords.), *Constituição Portuguesa Anotada*, 2.ª ed., Wolters Kluwer, Coimbra, 2010, p. 1312.

[26] Entende ainda a OMS que o impacto do etarismo verifica-se mais em grupos menos favorecidos, sendo que pessoas mais velhas com menos escolaridade são mais propensas a experienciar as consequências do etarismo no setor da saúde. Veja-se https://iris.who.int/bitstream/handle/10665/340208/9789240016866-eng.pdf?sequence=1&isAllowed=y, p. 2.

[27] Luísa Pinheiro, «Idadismo: a última discriminação socialmente aceite», in *O Direito dos "mais velhos"*, *E-book*, CEJ, 2019, p. 44, disponível em https://cej.justica.gov.pt/LinkClick.aspx?fileticket=AxgUgz4KFVc%3d&portalid=30 [consultado em 28/09/2025].

[28] *Ibidem*.

e metas a que se propõem – que, note-se, são muitos –, é necessário garantir e assegurar que os direitos em causa não são violados e desprezados.

Sendo o envelhecimento populacional um fator a considerar pelos sistemas inteligentes, estes tenderão a estabelecer onde priorizar os recursos para cuidados clínicos. Repare-se que, como já referido, os dados acumulados pela IA tendem a excluir grupos tendencialmente menos favorecidos ou minoritários. Não obstante, a discriminação com base na idade consubstanciará, certamente, uma violação do princípio da igualdade, nomeadamente pelo facto de a desigualdade de tratamento pretender desqualificar quem a sofre, pretendendo introduzir uma diferenciação indiciadora de distinta dignidade[29]. O Regulamento de IA – *AI Act* – prevê, por isso, no seu artigo 5.º, a proibição de práticas que explorem vulnerabilidades, nomeadamente em função da idade, e, com isso, originem riscos ou prejuízos físicos ou psicológicos.

É por isso essencial que o envelhecimento populacional deva ser um fator preponderante nas políticas de saúde, já que as condições de vulnerabilidade da pessoa idosa podem ser semelhantes à de uma criança ou jovem em perigo. Aliás, é também esse o entendimento da doutrina, quando refere que o direito fundamental à proteção na saúde é composto por duas vertentes: positiva e negativa. Por um lado, o direito a prestações por parte do Estado e, por outro lado, o direito a que o Estado e terceiros não se abstenham de prejudicar o bem jurídico "saúde"[30].

3. A (dependência da) IA no setor da saúde

Nos últimos anos e décadas, assistimos a uma crescente digitalização dos cuidados de saúde, desde serviços médicos, agendamento de consultas e exames a funções mais elaboradas como cirurgias ou diagnósticos como suporte para a decisão. A IA é, por isso, uma ferramenta importante para o desenvolvimento da saúde nos próximos anos. Esta imposição da IA está, de facto, a gerar mudanças acentuadas quer a nível legislativo, uma vez que os Estados têm vindo a desenvolver preocupações sobre esta matéria, quer a nível estrutural e social.

Atualmente, há diversos desafios no setor da saúde, desde logo a nível económico e social. O envelhecimento populacional constitui também ele um desafio, já que, não obstante o aumento do número de anos vividos, representa também mais investimentos no final de vida, o que certamente poderá tornar-se mais eficiente com recurso à IA. Com um raciocínio lógico-dedutivo poderá concluir-se que, por via de sistemas inteligentes e automatizados, os sistemas de

[29] Cláudia Monge, «A proteção e a promoção da saúde da pessoa idosa», *op. cit.*, p. 143.

[30] Rui Medeiros, «Artigo 64.º», *op. cit.*, pp. 1305-1322.

saúde podem oferecer mais cuidados com menos recursos, sendo importante destacar que a saúde digital deverá encontrar-se em conformidade com a demais legislação em vigor. Para tal, será também necessária a confiança nestes sistemas por parte da comunidade.

Associada a este fenómeno surge também a questão da responsabilidade face aos danos e erros provocados pelos sistemas de IA que iremos analisar *infra*.

3.1. *Os diagnósticos e as decisões tomadas por sistemas inteligentes: a questão da responsabilidade*

Conforme analisado, a evolução tecnológica ofereceu diversos mecanismos para concretizar os cuidados de saúde de forma mais eficiente, dando-se a generalização dos cuidados de saúde acessíveis por vias digitais. Não obstante, desta crescente digitalização podem surgir erros que, por sua vez, causem danos. Podemos analisar esta problemática de duas perspetivas: por um lado, uma crescente perda de perícia e competência dos profissionais pela crescente prática por robots – *deskilling* – e, por outro lado, os riscos associados à confiança excessiva nos sistemas de IA e nos robots – *overfaith*[31].

Note-se que estes sistemas, pelas suas vicissitudes, geram especiais dificuldades na averiguação da responsabilidade civil. Sabemos que, por regra, o médico não será obrigado a curar ou restabelecer o estado de saúde do paciente, prestando os cuidados necessários adaptados a cada caso – obrigação de meios[32]. Esta obrigação visa resolver o problema associado ao ónus de prova na responsabilidade civil médica[33]. Para apurar a responsabilidade civil por erro médico, pressupõe-se a prática de um facto que, por sua vez, seja ilícito e dependente de culpa, existindo um dano e um nexo de causalidade correlacionado.

No que respeita à IA, foi elaborada a Resolução do Parlamento Europeu de 20 de outubro de 2020, considerando não ser necessário rever os regimes de responsabilidade civil que funcionem bem, mas a multiplicidade de intervenientes – como o programador, o fabricante, o vendedor, o comprador ou utilizadores – constitui, de facto, um desafio à "eficácia das disposições do qua-

---

[31] Rute Teixeira Pedro, «Breves reflexões sobre a reparação de danos causados na prestação de cuidados de saúde com utilização de robots», in Eva Sónia Moreira da Silva, Pedro Miguel Freitas (coords.), *Inteligência Artificial e Robótica*, Gestlegal, Coimbra, 2022, pp. 158-159. Disponível em https://repositorium.uminho.pt/server/api/core/bitstreams/a4f4478c-92d6-4e38-85ba-a73cdd8215da/content [consultado em 28/09/2025].

[32] Além da obrigação principal de tratamento, surgem também associadas obrigações acessórias, como a segurança dos equipamentos ou o dever de informar. Veja-se Maria João Estorninho, Tiago Macieirinha, *Direito da Saúde*, *op. cit.*, p. 205.

[33] *Ibidem*, p. 203.

dro de responsabilidade da União e dos Estados-Membros", sendo necessários ajustamentos "específicos e coordenados aos regimes de responsabilidade para evitar situações em que as pessoas que sofrem danos ou prejuízos patrimoniais não recebem uma indemnização". No âmbito da prestação de cuidados de saúde, vigora o princípio geral estabelecido no n.º 2 do artigo 483.º do Código Civil, isto é, apenas nos casos previstos na lei surge, independentemente da culpa, a obrigação de indemnizar.

O conhecimento adquirido com base na autoaprendizagem, utilizando as informações obtidas para, a partir daí, proceder às decisões, acentua também as dificuldades associadas ao nexo de causalidade, evidenciando a dificuldade da tarefa probatória[34]. A reflexão que se pretende fazer prende-se também com a imprevisibilidade associada às decisões tomadas por estes sistemas, o que coloca em questão a ilicitude e a culpa do prestador de saúde. Nos termos do n.º 1 do artigo 483.º do Código Civil, a ilicitude traduz-se, em regra, na violação de um direito absoluto ou de uma norma destinada a proteger os interesses do doente[35].

A doutrina tem entendido que, neste âmbito, podemos convocar o previsto no artigo 493.º do Código Civil[36]. Trata-se, por um lado, da presunção no n.º 1 do referido artigo de que recai sobre quem tiver em seu poder coisa, móvel ou imóvel, com poder de a vigiar, a responsabilidade pelos danos que a coisa causar – "salvo se provar que nenhuma culpa houve da sua parte ou que os danos se teriam igualmente produzido ainda que não houvesse culpa sua". Por outro lado, nos termos do n.º 2, quem causar danos a outrem no exercício de uma atividade perigosa é obrigado a repará-los – "exceto se mostrar que empregou todas as providências exigidas pelas circunstâncias com o fim de os prevenir". Por norma, não serão os cuidados de saúde uma atividade perigosa, podendo, no entanto, ponderar-se tal qualificação por força da especial perigosidade dos sistemas e métodos utilizados. Nos termos da *supra* referida Resolução do Parlamento Europeu, o utilizador responderá objetivamente independentemente da culpa pelos danos produzidos, sendo que aqui se discute também a possibilidade de atribuição de personalidade eletrónica ao robot em razão da sua crescente autonomia e autoaprendizagem, poder de decisão e adaptação comportamental, sendo que, para a maioria da doutrina, tal caracterização não procede, já que, ainda que seja elevado o nível de sofisticação, tal baseia-se em combinações algorítmicas reflexo de comportamentos padronizados[37].

---

[34] Rute Teixeira Pedro, «Breves reflexões sobre a reparação de danos…», *op. cit*., pp. 158-159.

[35] *Ibidem*, p. 171.

[36] *Ibidem*, p. 172.

[37] Mafalda Miranda Barbosa, *Inteligência Artificial – Entre a Utopia e a Distopia*, *op. cit.*, p. 200.

A *supra* referida Resolução do Parlamento Europeu sobre o regime da responsabilidade civil aplicável à IA considera que o sucesso dos produtos e serviços e das tecnologias dependerá da cobertura da responsabilidade como garantia de confiança nos referidos sistemas, nomeadamente através de seguros e pagando o respetivo prémio.

4. Conclusões

Nos últimos anos e décadas, tem-se assistido a um crescimento e desenvolvimento ímpar da tecnologia e, em particular, da IA. Atualmente já parecem distantes os tempos em que era necessária a deslocação à unidade hospitalar simplesmente para marcar uma consulta. A inclusão – e ousamos dizer até mesmo dependência – das tecnologias no quotidiano tem criado diversos desafios a vários ramos do Direito. A par destes desafios surgem também os associados ao envelhecimento populacional e ao aumento da esperança média de vida – não fôssemos um dos países mais envelhecidos do mundo.

Se parece inegável que a intervenção de sistemas inteligentes e altamente tecnológicos contribui para uma maior rentabilização de recursos, podendo ser tomadas decisões mais rápidas e mais eficientes, a verdade é que, à medida que a tecnologia se desenvolve, vão surgindo soluções mais inovadoras e com diagnósticos mais precisos. Não obstante, também deverá haver consciencialização para a garantia da proteção na saúde e no acesso aos tratamentos adequados às necessidades de cada um, já que tal desenvolvimento se deve pela crescente utilização de dados e uso de tecnologias com poder de autoaprendizagem. Assim, deverá haver cuidados acentuados para a proteção dos referidos direitos, nomeadamente quando em causa se encontram os referidos mecanismos de inteligência. Como sabemos, as decisões tomadas baseadas em algoritmos nem sempre são imparciais e desprovidas de preconceito, o que certamente acarretará riscos associados ao etarismo e à discriminação das pessoas mais velhas.

TITLE
E.TEC YEARBOOK

EDITORS
Prof.ª Doutora Sónia Moreira

AUTHORS
Abner Cabral Buccini | Ana Cristina Gouveia | Ana Sirage Coimbra
Filipe Venade de Sousa | Hana Horak | Luís Manuel Pica | Ng Chan Fai
Pedro Henrique Engler Urso; Matheus Raphael Pova Marinho Dantas
Sofia Ferreira da Costa

DATE
December 2025

PUBLISHERS
School of Law - University of Minho (www.direito.uminho.pt)
JusGov - Research Centre for Justice and Governance (www.jusgov.uminho.pt)

ISSN
2184-707X

University of Minho
School of Law

JusGov
RESEARCH CENTRE FOR JUSTICE AND GOVERNANCE

fct Fundação para a Ciência e a Tecnologia