DIREITO EM MUDANÇA

A Proposta de Regulamento Europeu sobre Inteligência Artificial

ALGUMAS QUESTÕES JURÍDICAS

Coordenação

SUSANA AIRES DE SOUSA

O presente livro foi realizado no âmbito da actividade da Área de Investigação «Risco, Transparência e Litigiosidade», integrada no projecto «Desafios Sociais, Incerteza e Direito: Pluralidade | Vulnerabilidade | Indecidibilidade» do Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra (UIDB/04643/2020)

Edição
Instituto Jurídico
Faculdade de Direito da Universidade de Coimbra

Concepção Gráfica
Tipografia Lousanense, Lda.

Contactos
geral@ij.uc.pt
www.uc.pt/fduc/ij
Colégio da Trindade | 3000-018 Coimbra

e-ISBN
978-989-9075-50-4

DOI
https://doi.org/10.47907/DireitoemMudanca/2023/livro

Instituto Jurídico | Faculdade de Direito | Universidade de Coimbra

DIREITO EM MUDANÇA

A Proposta de Regulamento Europeu sobre Inteligência Artificial

ALGUMAS QUESTÕES JURÍDICAS

Coordenação

SUSANA AIRES DE SOUSA

1 2 9 0

INSTITVTO IVRIDICO
FACULDADE DE DIREITO
UNIVERSIDADE D
COIMBRA

Índice

Nota prévia

Na sua atividade de investigação, o Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra tem vindo a acolher e a promover o estudo e o debate em torno dos desafios que os sistemas de inteligência artificial (IA) colocam ao Direito. Um marco inevitável neste difícil equilíbrio, entre uma tecnologia aceleradamente dinâmica na sua evolução e a natureza tendencialmente estática das normas jurídicas, é a regulação destes sistemas, assumindo particular relevância a "Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de Inteligência Artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União". A importância desta Proposta justifica-se tanto pelo seu impacto no contexto europeu, como também pelos efeitos que pode ter na regulação dos sistemas de IA no plano internacional, do ocidente ao oriente.

Num esforço de estimular o debate em torno das soluções encontradas nesta Proposta de Regulamento, simultaneamente inovadora e de enorme atualidade, promoveu o Instituto Jurídico, no passado dia 28 de março, o *Webinar* "A proposta de Regulamento Europeu sobre Inteligência Artificial: Algumas Questões jurídicas". As apresentações, o debate e a intensa discussão propiciada por este encontro confirmaram a atualidade do tema e constituíram um forte impulso não só para dar forma escrita às exposições orais, como também para adicionar novas reflexões sobre aquela proposta regulatória da União Europeia. Esta obra corresponde à concretização deste propósito, impondo-se, de forma imediata, o necessário agradecimento aos autores dos vários capítulos por, através deles, generosamente partilharem o seu estudo e o seu saber sobre um documento ainda em construção no momento em que se escrevem estas linhas.

Esta publicação procura ainda responder a um outro desafio que o Instituto Jurídico, enquanto unidade de investigação, tem vindo a

assumir e que se concretiza no envolvimento de estudantes de doutoramento e de mestrado nas atividades desenvolvidas. Os jovens investigadores tomaram parte no momento de debate e nas reflexões que a ele se seguiram. Assim, cumprindo-se também este outro propósito, integra-se nesta obra um primeiro estudo do grupo de investigação "Inteligência Artificial e manutenção da Ordem Pública: impacto da proposta de regulamento de inteligência artificial no direito português", formado por estudantes de mestrado e de doutoramento a partir da iniciativa *Researchers' Camp*, promovida pelo Instituto Jurídico com a finalidade auxiliar jovens investigadores a progredir na carreira de investigação.

O *webinar* e a obra escrita que ora se publica enquadram-se na iniciativa "Direito em Mudança", que a área de investigação *Risco, Transparência e Litigiosidade* tem vindo, de forma contínua e persistente, a desenvolver nos últimos anos, sob o estímulo constante e a confiança inabalável da Senhora Doutora Maria João Antunes, enquanto Coordenadora desta área, a quem deixamos o nosso reconhecido agradecimento.

Susana Aires de Sousa
Em 2023, nos primeiros dias de junho.

1

Breves notas sobre a "Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de Inteligência Artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União"

(https://doi.org/10.47907/DireitoemMudanca/2023/1)

*Susana Aires de Sousa**

Resumo: Neste texto faz-se um breve enquadramento da Proposta de Regulamento Inteligência Artificial, estruturado e desenvolvido em quatro breves notas: feita uma introdução ao tema, faz-se uma contextualização daquele documento legislativo, para, em seguida se abordar o regulamento propriamente dito nos seus elementos principais (objetivo e âmbito de aplicação, por um lado, e definição de sistema de IA e estrutura do documento, por outro lado). O texto termina com algumas considerações sobre as etapas legislativas seguintes.

Palavras-chave: Sistemas de IA, Regulação, Proposta de Regulamento União Europeia

1. Nota introdutória

Os algoritmos são omnipresentes nas nossas vidas e na realização dos nossos interesses, interagindo com humanos e outros algoritmos.

* Univ Coimbra, IJ, Faculdade de Direito, Portugal, ORCID 0000-0003-4808-2466

Esta presença algorítmica é, simultaneamente e quase sempre, pela sua natureza, invisível aos olhos humanos. Diluindo-se a fronteira entre inteligências (da máquina e do humano), os sistemas computadorizados complexos integram as decisões humanas quotidianas de muitos e diversos modos ainda que, por vezes, ausentes à imediata consciência humana: em indicações ou informações (meteorológica, gastronómica, geográfica, etc.) prestadas pelo *smartphone*; na geolocalização; na assistência virtual a encomendas, transporte e entrega de bens ou prestações de serviços; na divulgação e gestão de publicidade adequada ao perfil pessoal; no jogo *online*; na concessão de crédito bancário; no sistema de orientação e de condução do veículo; na enorme precisão do robô cirúrgico utilizado na intervenção cirúrgica; em *trading* de ações e produtos financeiros; na avaliação da mais-valia laboral de um trabalhador, entre muitos outros exemplos.

Esta relação próxima, de quase dependência, entre o humano e o algoritmo fez-se merecedora de atenção à medida em que, no contexto daquela relação, a máquina vai ganhando autonomia decisória, através da sua capacidade para aprender. Em vez de calcular a melhor opção de entre aquelas que integram uma base de dados históricos, o sistema consegue, comparar padrões e reconstruir o modelo de forma a chegar a um resultado novo, mais eficiente em face do objetivo proposto. Contudo, esta natureza dinâmica e disruptiva da máquina tornou visível alguns dos riscos ligados à sua imprevisibilidade, velocidade e incontrolabilidade. Os riscos inerentes à IA saíram da esfera do programador e do *provider* e ganharam consciência coletiva na medida em que deixaram de ser vislumbrados como um mero acaso, mas como fontes de possível responsabilidade. O conflito entre inovação e precaução ganha, sobretudo na última década, enorme visibilidade. Ora, a regulação normativa da IA impõe-se quer como instrumento de gestão e controlo dos riscos inerentes aos sistemas computadorizados complexos, quer ainda como poderosa ferramenta na distribuição de responsabilidade por possíveis danos ligados àqueles sistemas.

É justamente o tema da regulação normativa da IA que nos ocupará. Em particular, nesta páginas e em jeito de introdução aos temas que se lhes seguem, procurar-se-á fazer um breve enquadramento da Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de Inteligência Artificial (Regulamento Inteligência Artificial) e altera determinados atos

legislativos da União, apresentada, a 21 de abril de 2021, pela Comissão Europeia, no contexto de uma estratégia europeia para a IA[1].

Desde esse momento, a proposta inicial tem vindo a ser debatida no Conselho e no Parlamento Europeu. Esse debate resultou já em várias modificações ao texto inicial, sobretudo depois da emergência e do enorme impacto da chamada "IA generativa"[2], capaz de criar conteúdos até aí inexistentes a partir dos dados que a alimentam.

Em traços largos, no texto que se segue, faz-se um breve enquadramento desta Proposta de Regulamento, estruturado em quatro breves notas: uma primeira de contextualização, seguida por duas breves notas sobre o Regulamento propriamente dito (objetivo e âmbito de aplicação, por um lado, definição de IA e estrutura da proposta, por outro lado), terminando com considerações sobre os momentos legislativos que se seguem.

2. Nota de contextualização

Não constitui qualquer novidade afirmar que, na última década, se colocou em evidência o impacto social, político e económico da Inteligência Artificial. Essa consciencialização da enorme capacidade *transformativa* da IA[3] provocou a necessidade da sua regulação. No

---

[1] A versão originária da proposta está disponível em https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

[2] Em causa estão modelos de aprendizagem profunda "capazes de gerar conteúdos muito semelhantes a resultados humanos (por exemplo, textos e imagens, palavras) em resposta a pedidos, questões, ou instruções que lhe são dirigidas", cf. Weng Marc Lim /Asanka Gunasekara / Jessica Leigh Pallant / Jason Ian Pallant / Ekaterina Pechenkina, «Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators«, *The International Journal of Management Education* 21/2 (2023) 100790. Exemplo deste tipo de tecnologia são o ChatGPT ou o Midjourney. Logo nos seus primeiros dias, o ChatGPT alcançaria milhões de utilizadores, tornando-se na aplicação tecnológica que mais rapidamente cresceu no espaço de poucos meses, cf. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

[3] Sobre esta natureza transformativa da IA em termos sociais e jurídicos veja-se Christoph Burchard, «Artificial Intelligence as the End of Criminal Law? On the Algorithmic Transformation of Society», in Maria João Antunes; Susana Aires de Sousa, coord., *Artificial Intelligence in the Economic Sector: prevention and responsibility*, Coimbra: Instituto Jurídico, 2021, 165 e ss., disponível em https://estudogeral.uc.pt/bitstream/10316/99251/1/Livro%20completo%20Artificial%20Intelligence%20%28online%29.pdf

espaço europeu, foi-se progressivamente afirmando esta necessidade de criar respostas regulativas, capazes de assegurar um desenvolvimento tecnológico integrado no respeito pelos valores da União e da Carta dos Direitos Fundamentais.

Essa preocupação ética com o desenvolvimento da IA foi logo avançada nos primeiros documentos[4] que emergiram no espaço europeu, como a *Comunicação da Comissão sobre Inteligência Artificial para a Europa* (também conhecida como Estratégia IA), de abril de 2018, e depois em outros documentos como a *Comunicação da Comissão sobre Aumentar a Confiança numa IA centrada no ser humano*, em 2019, ou as Orientações Éticas para uma IA de Confiança, publicadas em Abril de 2019 pelo Grupo Independente de Peritos de Alto Nível, criado pela Comissão Europeia. O *White Paper* sobre IA, também da Comissão, de fevereiro de 2020 constituiu um marco importante na história da regulação europeia da IA na Europa, uma vez que nele se desenha uma aproximação normativa aos sistemas de IA a partir do risco que lhe é inerente. Muitos outros documentos, entre Comunicações, Relatórios e Resoluções, foram surgindo, no contexto europeu, entre os quais cabe destacar, as várias Resoluções adotadas pelo Parlamento Europeu, em outubro de 2020, ligadas aos problemas éticos da IA, à responsabilidade civil por IA ou à propriedade intelectual, solicitando ainda à Comissão que estabeleça um quadro legal para o desenvolvimento, disponibilização e uso de sistemas de IA, sistemas robotizados e tecnologias semelhantes.

É justamente nesta conjuntura regulativa, que se esperava ser longa, que a Comissão Europeia torna pública uma proposta de Regulamento sobre a IA (*AI Act*), em 21 de abril de 2021, que, entretanto, e desde então, tem vindo a ser objeto de ampla discussão, no contexto académico e também entre as instituições europeias. Sobre esta proposta emitiram parecer positivo, em alguns casos com recomendações de

---

[4] Os documentos referidos bem como outras fontes relevantes podem encontrar-se na compilação realizada ao abrigo do Projeto Exploratório IA e Criminalidade Empresarial, do Instituto Jurídico, e estão disponíveis em https://www.uc.pt/fduc/ij/projetos-de-investigacao/inteligencia-artificial-e-criminalidade-empresarial/regulacao-normativa-da-inteligencia-artificial/. Veja-se ainda, sobre estes documentos, José Ricardo Marcondes Ramos, «Relatório sobre a atual regulação normativa europeia e portuguesa em matéria de Inteligência Artificial», *Revista Portuguesa de Ciência Criminal* 31/3 (2021) 633-646.

alteração, o Comité Económico e Social Europeu[5], em 22 de setembro de 2021, o Banco Central Europeu, em 29 de dezembro de 2021[6], e o Comité Europeu das Regiões, em 2 de dezembro de 2021[7]. Simultaneamente, a 23 de abril de 2021, iniciavam-se os debates e as discussões no contexto do Conselho Europeu, sendo apresentada e pulicada uma nova versão da proposta, a 6 de dezembro de 2022, enunciadora da posição do Conselho e que constituiria o texto-base das negociações entre os grupos parlamentares do Parlamento Europeu.

Este texto, com alterações, emendas e ajustes resultantes dos diversos debates que foram sendo realizados, é, no momento em que se escreve este texto, objeto de análise, tendo em vista a obtenção de uma concordância entre grupos parlamentares. As notícias mais recentes dão conta de algumas modificações sugeridas por esta discussão prévia entre os grupos parlamentares, designadamente quanto ao conceito de inteligência artificial e quanto ao enquadramento jurídico da chamada "*generative AI*".

A análise que se segue toma por objeto a versão mais recente da "Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da União", divulgada em 6 dezembro de 2022.

3. O Regulamento (primeira nota): objetivo e âmbito de aplicação

Com esta Proposta pretende-se que uma tecnologia em acelerada expansão seja enquadrada na observância de valores éticos e jurídicos estruturantes da União Europeia, adivinhando-se, desde o início que este seria um processo legislativo longo, desde logo, pelas dificuldades em abranger e aprender do ponto de vista normativo uma tecnologia de natureza disruptiva e imprevisível.

---

[5] O documento está disponível em https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52021AE2482

[6] Disponível em https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52021AB0040

[7] Disponível em https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52021AR2682

Contudo, é importante sublinhar que este esforço legislativo da União Europeia constitui a primeira tentativa de criar uma regulamentação geral e uniforme para a IA e, como tal, um marco a nível internacional, que procura assegurar que, no mercado europeu, esta tecnologia se desenvolverá de forma eticamente fundada, socialmente justa e equitativa e ambientalmente sustentável[8]. Constituindo uma das preocupações da proposta estabelecer "um quadro jurídico uniforme para o desenvolvimento, a comercialização e a utilização da inteligência artificial" que evite a adoção de regras nacionais muito díspares em prejuízo da unidade do mercado interno, da segurança jurídica e dos valores éticos que pautam a União Europeia, compreende-se a opção por um Regulamento e não por qualquer outra forma legislativa que implicasse a adaptação ou transposição pelos Estados-membros. Deste modo, o Regulamento, sendo de aplicação geral, não carece de qualquer transposição pelos Estados-membros, entrando em vigor, em todos o Estados na data fixada, à semelhança do que aconteceu, por exemplo, com o Regulamento Geral de Proteção de Dados.

Assim, o Regulamento, comum ao espaço da União Europeia (e por tanto às entidades públicas e privadas que atuam no espaço europeu), procura ser uma proposta equilibrada, entre, por um lado, o respeito aos direitos fundamentais e, por outro, o estímulo do desenvolvimento tecnológico, fundamental para uma europa economicamente forte capaz de concorrer a oriente e a ocidente. O desígnio essencial desta proposta é criar um regime jurídico para a IA que garante que os produtos colocados e utilizados no mercado da UE são seguros e confiáveis, mas também conformes aos direitos fundamentais e valores da União. Ou seja, nas palavras da Proposta, promover o desenvolvimento de novas tecnologias, assegurando um nível elevado de "proteção de interesses públicos, como a saúde e a segurança e a proteção dos direitos fundamentais, conforme reconhecido e protegido pelo direito da União".

Este objetivo geral pode, por sua vez, ser decomposto em quatro linhas principais: assegurar que os sistemas de IA respeitam os valores éticos da EU e a legislação em vigor sobre direitos fundamentais; garantir segurança jurídica por forma a atrair investimento na área tecnológica; promover uma boa governação que permita uma IA segura

---

[8] Luciano Floridi, «The European Legislation on AI. A Brief Analysis of its Philosophical Approach», *Philosophy & Technology* 34 (2021) 215-222.

e conforme aos direitos fundamentais; promover um mercado único e uniforme para os sistemas e aplicações de IA.

Cada uma destas linhas é materializada ao longo dos artigos que integram a proposta de Regulamento, designadamente em deveres e obrigações a cumprir por entidades que produzam, disponibilizem, importem ou utilizem determinados sistemas de IA.

4. O Regulamento (segunda nota): definição de IA e estrutura da Proposta

O conceito de Inteligência Artificial tem sido objeto de enorme discussão, na medida em que procura conjugar dois elementos de difícil conciliação: por um lado, encontrar um conceito preciso e inequívoco de modo a assegurar a segurança jurídica, mas, por outro lado, ser suficientemente flexível para se adaptar a futuras evoluções tecnológicas[9]. Compreende-se, assim, que a definição constante da proposta original tenha vindo a sofrer várias modificações e emendas. As notícias sobre a discussão em torno do conceito apontam para que no Regulamento venha a seguir-se um conceito de sistema IA próximo daquilo que tem vindo a ser defendido pela OCDE, entendido como "um sistema de aprendizagem concebido para atuar com diversos níveis de autonomia, capaz de gerar, para determinados fins, explícitos ou implícitos, resultados como predições, recomendações ou decisões aptas a influenciar ambientes físicos ou virtuais"[10].

O Regulamento faz uma abordagem e uma classificação dos sistemas de IA a partir do seu grau de risco para direitos fundamentais. Contudo, há, à partida, sistemas que, independentemente do seu grau de risco, ficam fora do âmbito de aplicação do Regulamento: é o caso sistemas para fins militares, de defesa ou de segurança nacional, independentemente do tipo de entidade, pública ou privada, que realiza essas atividades[11]; é também o caso de sistemas desenvolvidos e

---

[9] Cf. Considerando 6 da Proposta de Regulamento (versão de 6 de dezembro de 2022).

[10] Cf., com indicações dos mais recentes movimentos legislativos no contexto do Parlamento Europeu, https://www.euractiv.com/section/artificial-intelligence/news/ai-act-moves-ahead-in-eu-parliament-with-key-committee-vote/

[11] Sobre este ponto, veja-se o texto, neste livro, de Miguel João Costa, «Sistemas de armas autónomas e respetiva regulação».

e utilizados exclusivamente para fins de investigação e desenvolvimento científicos, de modo a assegurar que a regulação não afete a atividade de investigação e desenvolvimento científicos em matéria de sistemas de IA.

A classificação dos sistemas atendendo ao seu grau de risco é gradativa e assenta em quatro categorias: *risco inaceitável*, por isso, tais sistemas são proibidos; sistemas de *risco elevado*; *risco limitado* e sistemas de *risco mínimo ou sem risco*. Esta abordagem a partir do grau de risco tem por objetivo principal proibir os sistemas de risco inaceitável e estabelecer requisitos específicos para sistemas de IA de risco elevado, bem como obrigações e deveres para os operadores destes sistemas (artigo 1.º). Deste modo, o quadro jurídico de obrigações será mais intenso conforme a gravidade de risco atribuída ao sistema.

Em termos de organização e de estrutura, a Proposta de Regulamento divide-se em vários Títulos e é completada por vários Anexos.

O primeiro Título é composto por disposições gerais sobre o âmbito de aplicação, o objetivo e as definições. Um novo título (I-A), dedicado aos sistemas de IA de finalidade geral, foi acrescentado à proposta originária.

Segue-se um conjunto de artigos que identifica as práticas de IA proibidas (Título II).

O Título III, dedicado aos sistemas de risco elevado, é o título mais extenso, com diversos capítulos, constituindo uma das parte mais relevantes deste diploma. É neste título que se preveem as regras de classificação de um sistema como de risco elevado[12], mas também, os requisitos e deveres a que ficam sujeitos este tipo de sistemas em matéria de gestão do risco, governação, documentação técnica e registo. É também nestes artigos que se encontra prevista a obrigação de identificar e analisar "riscos conhecidos e previsíveis mais suscetíveis de ocorrer para a saúde, a segurança e os direitos fundamentais" ou de adotar e manter, durante o ciclo de vida do sistema de IA, um plano de gestão e de avaliação de riscos, ou ainda, as regras referentes à documentação técnica, as obrigações de transparência, de informação e de registo a

---

[12] As últimas notícias dão conta de que no debate sobre a Proposta ocorrido no Parlamento Europeu se terá definitivamente afastado, como técnica legislativa, uma enunciação taxativa dos sistemas de elevado risco, com recurso a uma cláusula aberta segundo a qual um sistema pode ser assim categorizado a partir do seu potencial risco para a saúde, segurança ou direitos fundamentais da pessoa.

que está sujeito um sistema que seja qualificado como de elevado risco. Ainda de notar, neste título, no capítulo 4, a obrigação de os Estados-membros de designar ou criar pelo menos uma autoridade notificadora responsável por estabelecer e executar os procedimentos necessários para a avaliação, a designação e a notificação de organismos de avaliação da conformidade do sistema de risco elevado. O capítulo seguinte deste título diz respeito às normas técnicas (*standards*), declaração de conformidade e certificação dos sistemas de elevado risco[13].

O Título IV compreende as obrigações de transparência aplicáveis aos fornecedores e utilizadores de determinados sistemas de inteligência artificiais, por exemplo, sistemas que se destinem a interagir com pessoas, sistemas de identificação biométrica, sistemas que visam o reconhecimento de emoções.

No Título V estipulam-se medidas de apoio à inovação, regulando-se por exemplo, a testagem em ambiente real de forma controlada.

O Título VI é dedicado à Governação[14]. Dedicam-se vários artigos à criação de um Comité Europeu para a Inteligência Artificial, aos deveres e poderes da Comissão e às Autoridades Nacionais competentes em matéria de IA.

A criação, pela Comissão e com a colaboração dos Estados-membros, de uma base de dados europeia relativa aos sistemas de inteligência artificial de risco elevado está regulada no Título VII.

O Título VIII compreende as medidas de acompanhamento e vigilância dos sistemas após a sua comercialização, a partilha de informações e a fiscalização do mercado.

A criação de códigos de conduta nesta matéria também não foi esquecida pelo Regulamento que expressamente se refere a este tipo de medidas no Título IX.

Os Estados membros devem ainda prever um regime sancionatório que previna de forma adequada a infração dos deveres e obrigações decorrentes do Regulamento, indicando-se expressamente as coimas aplicáveis à violação de determinados deveres no Título X.

---

[13] Em geral, com referência a estas obrigações, veja-se, neste livro, o texto de José Ricardo Marcondes Ramos, «Supervisão, classificação e certificação dos sistemas de IA na Proposta de Regulamento sobre inteligência artificial».

[14] Em ligação com este ponto, neste livro, veja-se o texto de Maria Elisabete Ramos, «Governação empresarial e gestão de risco de IA».

O Título XI refere-se aos poderes delegados na Comissão e à possibilidade de esta ser assistida por um Comité.

Por fim, prevê-se um último título com disposições finais.

Também os Anexos foram sofrendo modificações entre a proposta originária e a versão publicada em dezembro de 2022. Desde logo, desapareceu o primeiro anexo. O segundo Anexo prevê os diplomas a harmonizar com o novo Regulamento. O terceiro Anexo, de grande relevância, faz um elenco de sistemas considerados como de risco elevado nos termos do artigo 6.º[15]. Segue-se o Anexo sobre a documentação técnica exigida nos termos do artigo 11.º do Regulamento. O Anexo V prevê a chamada declaração de conformidade a que se refere o artigo 48.º do Regulamento. O Anexo VI refere-se ao procedimento de avaliação da conformidade a partir do controlo interno (exigido pelo artigo 17.º do Regulamento). O Anexo VII contempla a conformidade baseada na avaliação do sistema de gestão de qualidade e na avaliação da documentação técnica (de acordo com o artigo 17.º do Regulamento). Por sua vez, o Anexo VIII diz respeito às informações a apresentar para fins de registo de operadores e de sistemas de Inteligência Artificial de risco elevado nos termos do artigo 51.º. Foi ainda adicionado um Anexo VIII-A relativo às informações a apresentar aquando do registo dos sistemas de IA de risco elevado elencados no Anexo III em relação à testagem em condições reais de acordo com o artigo 54.º-A. Por fim, o Anexo IX refere-se à legislação europeia relativa a sistemas informáticos de grande escala no espaço de liberdade, segurança e justiça.

Desta breve enunciação da organização sistemática do Regulamento, sujeita ainda a alterações que decorram do Parlamento e do Conselho, evidencia-se uma intenção com deveres de transparência, de gestão de qualidade, de governação de dados, de gestão de risco, de supervisão humana, de robustez do sistema em matéria de cibersegurança. É dada particular atenção aos sistemas qualificáveis pela Comissão como de elevado risco (art. 7.º) (por exemplo, sistemas em que haja risco de ofensa à saúde e segurança, em contexto médico ou terapêutico), ou risco elevado e muito provável de ofensa aos

---

[15] Com referência a este anexo e, em particular, aos sistemas de IA para fins de manutenção da ordem pública, neste livro, o texto de Alberto Raphael Ribeiro Magalhães / Ana Cristina Crestani / Luiza Tosta Cardoso Franco, «Inteligência artificial no âmbito da manutenção da ordem pública: considerações iniciais sob a ótica da Proposta de Regulamento do Parlamento Europeu e do Conselho».

direitos fundamentais) ou ainda os sistemas listados como de elevado risco no Anexo III do Regulamento (identificação biométrica, gestão e funcionamento de infraestruturas críticas, saúde, educação, manutenção da ordem pública). Estes sistemas terão de facto de estar em *compliance* com as obrigações de certificação, validação, supervisão, governação e transparência impostas pelo Regulamento.

Sublinhe-se que, não obstante os sistemas de elevado risco concentrarem grande parte dos artigos que integram o Regulamento, há obrigações que se estendem aos sistemas de risco limitado. É o caso, por exemplo, da exigência de transparência. Por sua vez, sistemas com risco mínimo não estarão sujeitos aos deveres impostos pelo Regulamento.

Como se enunciou, há sistemas que se quis expressamente proibir no espaço europeu (elencados no artigo 5.º), como, por exemplo, a utilização destes sistemas para fins de *ranking*, pontuação ou classificação social ou que explorem vulnerabilidades de um grupo específico de pessoas. Outros que, partindo da sua não admissibilidade, podem contudo vir a ser admitidos, como a utilização de identificação biométrica, em princípio proibida, em espaços públicos e em tempo real, mas admissível, mediante autorização, para efeitos de manutenção da ordem pública[16].

Os textos que se seguem exploram de forma mais substancial e fundamentada algumas das ideias agora exposta de forma e sumária, seja no contexto da governação das empresas, da saúde, das armas autónomas e da manutenção da ordem pública.

5. Nota sobre o itinerário legislativo futuro

Em dezembro de 2022, o Conselho aprovou uma versão de compromisso que reflete uma posição comum e publicou uma versão da Proposta de Regulamento que integrava as alterações propostas.

Os dois relatores Parlamento Europeu, Dragos Tudorache e Brando Benifei, comunicaram, em fevereiro de 2023, os pontos fundamentais

---

[16] Sobre este ponto, Rui Soares Pereira, «Sobre o uso de sistemas de identificação biométrica (e de tecnologias de reconhecimento facial) para fins de segurança pública e de aplicação coercitiva da lei: reflexões a propósito da proposta de regulamento europeu sobre a inteligência artificial», *Revista da Faculdade de Direito da Universidade de Lisboa* 63 (2022) 839 e ss.

que deveriam ser objeto de consenso, com ênfase na definição de IA, na delimitação das práticas proibidas e na indicação de critérios que facilitem a qualificação de sistemas de IA como de risco elevado. Já em março soube-se que a definição de IA é um dos pontos mais críticos que provavelmente sofrerá ainda alterações. Seguiu-se a discussão no Parlamento Europeu, tendo os relatores do projeto promovido várias reuniões com os vários grupos políticos. Estes debates permitiram algum consenso, mas deram origem a novas emendas publicadas em 9 de maio de 2023[17].

De facto, no início de maio[18], o *Internal Market and Civil Liberties Committees*, do Parlamento Europeu, aprovou a uma versão de consenso da Proposta, com 84 votos a favor, 7 votos contra e 12 abstenções[19]. Como é sublinhado na comunicação pública sobre esta negociação, alguns dos pilares desta proposta referem-se a uma definição de sistema de IA que possa abranger não só os sistemas já existentes como aqueles que venham a ser desenvolvidos, atendendo à enorme velocidade do desenvolvimento tecnológico. Houve ainda consenso quanto à proibição de sistemas de IA com fins de manipulação ou que explorem vulnerabilidade ou que possam ser usados para fins de pontuação social (*social scoring*), bem como a aplicações intrusivas ou discriminatórias, incluindo sistemas de identificação biométrica remota, em tempo real, em espaços acessíveis ao público. Ampliou-se a lista de sistemas classificáveis como de risco elevado, de forma a abranger sistemas perigosos à saúde, segurança, direitos fundamentais e ambiente. Há ainda outras modificações relevantes, algumas das quais procuram promover a inovação ao excluir das obrigações previstas a atividade de investigação ou ao dar prioridade a uma regulação que privilegie a *sandbox aproach*.

As últimas notícias dão também conta da discussão sobre a inclusão no Regulamento da chamada *generative AI* – a IA com capacidade

---

[17] Documento disponível em https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf?fbclid=IwAR3dIussPXxnMPQgTey9_a4tpSdZxFjTuF1nYWLK2xZdgoA1BFJstjJRj40

[18] https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence?utm_source=substack&utm_medium=email

[19] Toda a informação e documentação sobre esta etapa pode encontrar-se no seguinte endereço: https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence

de gerar novos conteúdos e com aplicação generalizada ou ampla, em múltiplos contextos. Os dois relatores europeus, Dragoș Tudorache e Brando Benifei, divulgaram no dia 14 de março uma proposta sobre este tema propondo algumas obrigações para aqueles que disponibilizam este tipo de sistemas. Por exemplo, propõe-se a introdução de um artigo que submete estes sistemas que produzem textos ou imagens criados pela máquina e que possam ser confundidos ou passar por conteúdo criado por humano, às mesmas obrigações de governação e de transparência a que estão sujeitos os sistemas de risco elevado, a menos que se identifique alguém como legalmente responsável por aquele conteúdo.

As negociações parciais no Parlamento Europeu chegaram ao fim no mês de maio, seguindo-se, em junho, a discussão e aprovação pelo plenário. Depois, o itinerário legislativo continuará por via das negociações no âmbito Conselho. Muito embora o percurso já efetuado evidencie passos largos e sólidos no sentido da adoção do Regulamento Europeu da IA, há certamente etapas por cumprir que levarão tempo. Entretanto, a necessidade de se conhecer a proposta legislativa, e, sobretudo, de refletir sobre as obrigações e deveres jurídicos nela contidos justificam os estudos que nesta obra se apresentam e que se seguem a este breve enquadramento.

Referências bibliográficas

Burchard, Christoph, «Artificial Intelligence as the End of Criminal Law? On the Algorithmic Transformation of Society», in Maria João Antunes; Susana Aires de Sousa, coord., *Artificial Intelligence in the Economic Sector: prevention and responsibility*, Coimbra: Instituto Jurídico, 2021, 165-200, disponível em https://estudogeral.uc.pt/bitstream/10316/99251/1/Livro%20completo%20Artificial%20Intelligence%20%28online%29.pdf

Floridi, Luciano, «The European Legislation on AI. A Brief Analysis of its Philosophical Approach», *Philosophy & Technology* 34 (2021) 215-222.

Lim, Weng Marc / Gunasekara, Asanka / Pallant, Jessica Leigh / Pallant, Jason Ian / Pechenkina, Ekaterina, «Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators«, *The International Journal of Management Education*, Vol: 21, Issue: 2, 2023.

Pereira, Rui Soares, «Sobre o uso de sistemas de identificação biométrica (e de tecnologias de reconhecimento facial) para fins de segurança pública e de aplicação coercitiva da lei: reflexões a propósito da proposta de regulamento europeu sobre a inteligência artificial», *Revista da Faculdade de Direito da Universidade de Lisboa* 63 (2022) 839-865.

Ramos, José Ricardo Marcondes, «Relatório sobre a atual regulação normativa europeia e portuguesa em matéria de Inteligência Artificial», *Revista Portuguesa de Ciência Criminal* 31/3 (2021) 633-646.

2

Governação empresarial e gestão de risco de IA

(https://doi.org/10.47907/DireitoemMudanca/2023/2)

Maria Elisabete Ramos[*/1]

Resumo: O presente trabalho trata, essencialmente, duas questões. A primeira consiste em saber que critérios devem orientar a decisão de utilizar sistemas de Inteligência Artificial (IA) como auxiliares do funcionamento do órgão de administração de sociedades. A segunda questão procura averiguar se o direito societário português admite ou não decisões dos administradores e deliberações dos sócios que procurem substituir a composição humana dos órgãos de administração

* Univ. Coimbra, CeBER, Faculty of Economics, Av. Dias da Silva 165, 3004-512 Coimbra. Professora Auxiliar com Agregação em Direito a exercer funções na Faculdade de Economia da Universidade de Coimbra. ORCID: 0000-0001-5376-4897.

[1] O presente texto corresponde à intervenção da sua Autora no *webinar Direito em Mudança. A proposta de Regulamento Europeu sobre Inteligência Artificial: algumas questões jurídicas*, organizada pela Senhora Profª Doutora Susana Aires de Sousa, Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, que teve lugar no dia 28 de março de 2023. O texto mantém o registo oral usado no *webinar*. Para mais informações (designadamente, bibliográficas), consultar Maria Elisabete Ramos, "CorpTech e desafios aos deveres de cuidado dos administradores", in *Inteligência artificial e robótica desafios para o direito do século XXI*, Eva Sónia Moreira da Silva (Sónia Moreira) e Pedro Miguel Freitas, coordenadores, Coimbra: Gestlegal, 2022, 229-250, disponível em https://www.jusgov.uminho.pt/pt-pt/publicacoes/inteligencia-artificial-e-robotica-desafios-para-o-direito-do-seculo-xxi/ (consulta no dia 4 de abril de 2023); Maria Elisabete Ramos; Ana Azevedo; Deolinda Meira; Mariana Curado Malta, «Cooperatives and the use of Artificial Intelligence: A critical view», *Sustainability*, 2023, 15, 329. https://doi.org/10.3390/su15010329, disponível em https://www.mdpi.com/journal/sustainability. Special Issue Co-operating for Change: Roles, Potentials, and Challenges of Cooperatives in the Decade Leading up to the Sustainable Development Goals.

por agentes de IA. A resposta à primeira questão convoca o "standard of conduct" dos deveres de cuidado dos administradores e a *business judgement rule* enquanto "standard of review" do dever de tomar decisões razoáveis. Os resultados da investigação motivada pela segunda questão mostram-nos que o direito societário português vigente não permite a substituição de órgãos societários "povoados por humanos" por agentes de IA.

Palavras-chave: Órgão de administração, deveres de cuidado, inteligência artificial, *business judgment rule*.

1. Dizem que a Inteligência Artificial é perfeita. A sério?

São repetidamente salientadas as vantagens do *desempenho da IA* relativamente à dos humanos. Tendo em conta a lista de vicissitudes – a que se vulgarizou designar "escândalos" – que ao longo da história têm evidenciado as falhas humanas na gestão das sociedades, a pergunta que se segue é se é na tecnologia que reside a solução para erradicar os históricos e nefastos conflitos de interesses e quebras de lealdade dos administradores. Para alguns, de facto, a tecnologia é a solução que, em última instância, irá permitir superar as imperfeições humanas, como são as atuações contaminadas por conflitos de interesses. E, desta forma, as sociedades, libertadas das imperfeições humanas e dos consequentes prejuízos, tornar-se-iam imaculados instrumentos de criação de riqueza e de prosperidade.

Alega-se que a tecnologia é imparcial, que não tem agenda própria e não está sujeita aos enviesamentos causados pelas emoções humanas. Objetividade e imparcialidade podem ser muito úteis ao eficaz funcionamento do órgão de administração, pois a tecnologia não é sensível às dinâmicas sociais que limitam, por exemplo, a expressão de dissidência. No entanto, os algoritmos de aprendizagem automática não são transparentes quanto ao *porquê* e ao *como* obtiveram determinado resultado, o que justifica a analogia de "black boxes". Circunstância que não travou a expansão destas tecnologias que, por um lado, são incomparavelmente eficazes na obtenção de resultados ou na formulação de previsões rigorosas e, por outro, são capazes de aprender e de evoluir, através da experiência e treino e sem supervisão humana.

Que as realizações humanas são imperfeitas, não necessita de demonstração; mas urge questionar a anunciada perfeição da IA, porque, de facto, estão identificadas várias limitações de que ela padece.

A primeira deve-se ao *enviesamento da informação* de que se serve a IA para produzir previsões ou recomendações. Se os dados a que a IA tem acesso e com que treina e aprende são enviesados ou de fraca qualidade – por exemplo, fundados em prévias decisões do conselho de administração que se mostraram erradas ou ilegais –, as decisões, recomendações ou previsões serão contaminadas por tais enviesamentos históricos.

Em segundo lugar, suscita-se o problema da *rastreabilidade da decisão*, pois as redes neurais profundas têm milhões de ligações que, no seu conjunto, formam a decisão tomada pela IA, mas é impossível rastrear o processo de tomada de decisão, como é difícil verificar se o algoritmo funciona corretamente. Os problemas de rastreabilidade determinam que seja difícil ou eventualmente impossível detetar e corrigir possíveis erros.

Sabe-se hoje que a IA é opaca na construção dos padrões, que a sua utilização representa riscos de violação de direitos humanos, discriminação algorítmica, intrusão na privacidade dos cidadãos, recolha ilegal de dados pessoais, manipulação de decisões e opiniões pessoais, violação de direitos dos consumidores, colusão algorítmica (Ezrachi e Stucke), através de *software* algorítmico de fixação de preços, finalidades criminais (ex. ciberataques)[2]. Não surpreende, pois, que a Proposta de Regulamento sobre IA eleja uma "abordagem baseada no risco"[3].

Muito recentemente, no âmbito da organização *Future of Life Institute*, foi lançada a iniciativa "Pause Giant AI Experiments: An Open Letter" em que os signatários "call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4."[4] Nesta carta lê-se: "Should we let machines flood our information channels with propaganda and untruth? Should we automate away all the jobs, including the fulfilling ones? Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete

---

[2] Sobre os riscos da IA, v. https://futureoflife.org/ai/benefits-risks-of-artificial-intelligence/ (consultado no dia 4 de abril de 2023).

[3] Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União {sec(2021) 167 final} - {swd(2021) 84 final} - {swd(2021) 85 final}, Bruxelas, 21.4.2021, COM(2021) 206 final, 7.

[4] Disponível em https://futureoflife.org/open-letter/pause-giant-ai-experiments/ (consulta no dia 4 de abril de 2023).

and replace us? Should we risk loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders. Powerful AI systems should be developed only once we are confident that their effects will be positive and their risks will be manageable."[5]

Da mole de problemas que esta *revolução tecnológica* traz ao direito societário, elegemos algumas questões suscitadas pela introdução de IA no processo de decisão do órgão de administração e os consequentes desafios aos deveres de cuidado dos administradores (humanos).

2. *CorpTech* – a IA auxilia ao funcionamento do *board*

O signo "CorpTech"[6] – abreviatura de *Corporate Technologies* – designa compreensivamente um *conjunto de tecnologias* como *big data analytics*, inteligência artificial, aprendizagem automática, *blockchain* e *smart contracts* aplicadas a matérias de *governo das sociedades* (latamente considerado) como a remuneração dos gestores, identificação de candidatos a posições cimeiras na organização, relação com investidores, o voto e os trabalhos do órgão de administração, gestão do risco, *compliance*.

Não estamos a falar do futuro, mas sim do presente. No entanto, o Código das Sociedades Comerciais não apresenta normas especificamente destinadas a regular a intervenção de sistemas de IA no processo de decisão do órgão de administração das sociedades, designadamente, não é identificado o órgão competente para tal decisão.

O art. 64.º, n.º 1, al. a), do Código das Sociedades Comerciais (CSC) consagra "deveres legais gerais"[7] de cuidado que, além de outros aspetos, exigem aos administradores "competência técnica e o conhecimento da atividade da sociedade adequados às suas funções". Por outro lado, as regras jurídico-societárias relativas à composição do órgão

---

[5] Disponível em https://futureoflife.org/open-letter/pause-giant-ai-experiments/ (consulta no dia 4 de abril de 2023).

[6] Luca Enriques / Dirk A. Zetzsche, «*Corporate Technologies and the Tech Nirvana Fallacy*», Working paper n.º 457/2019, March 2020, 4, [em linha] [consulta em 12/3/2022]. Disponível em https://ecgi.global/sites/default/files/working_papers/documents/finalenriqueszetzsche.pdf

[7] Cfr J. M. Coutinho de Abreu, «Deveres de cuidado e de lealdade dos administradores e interesse social», *Reformas do Código das Sociedades*, Coimbra: Almedina, 2007, 17 (15-47).

de administração não exigem requisitos de literacia em IA. Como se pode atuar em termos informados, se não se possui a suficiente *expertise* em matérias de IA? E se *a posteriori* a decisão se mostra errada e danosa para a sociedade e terceiros, porque, por exemplo, ocorre o "algorithm failure"? Estão as normas relativas aos deveres de cuidado suficientemente preparadas para enfrentar/gerir os desafios postos pela introdução da IA no processo de decisão do órgão de administração? E a opacidade do processo seguido pela IA na construção de certa recomendação ou previsão é compatível com os deveres gerais de cuidado que assumem uma manifesta dimensão procedimental?

O atual quadro legal regulador das sociedades comerciais e civis em forma comercial torna obrigatório, em todos os tipos societários, o órgão de administração e de representação (arts. 191.º, 252.º, 405.º e 470.º do CSC). Embora o vigente rol de competências do órgão de administração e de representação da sociedade seja sensível ao *tipo societário*, o modelo legal assenta no pressuposto de que as decisões deste órgão (imputadas juridicamente à sociedade) são tomadas por *humanos*.

Não é facto novo que o órgão de administração se sirva de programas de computador para preparar as suas decisões. Sublinha o considerando (6) da Proposta de Regulamento IA que uma das principais características funcionais do *software de IA* consiste na "capacidade, tendo em vista um determinado conjunto de objetivos definidos pelos seres humanos, de criar resultados, tais como conteúdos, previsões, recomendações ou decisões que influenciam o ambiente com o qual o sistema interage, quer numa dimensão física, quer digital"[8].

Face a esta novidade consistente no *desempenho* da IA, quais são as orientações que o órgão de administração "povoado por humanos"[9] poderá extrair dos deveres de cuidado? Aceitar a decisão/recomendação preparada pela IA, confiando nela, sem escrutínio ou, ao invés, assegurar-se que a decisão cumpre a lei (por exemplo, que não implica discriminação em razão do género ou etnia ou que cumpre as regras da concorrência)? Imagine-se, por exemplo, que a IA recomenda a celebração de um acordo com uma empresa concorrente que, de facto,

---

[8] Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial), considerando (6).

[9] A versão inglesa desta expressão pertence a Luca Enriques/Dirk A. Zetzsche, *Corporate Technologies and Tech Nirvana Fallacy,* 8.

configura um cartel de preços ou de repartição de mercados ou um cartel *hub and spoke*[10].

Parece-nos que compete ao órgão de administração e de representação decidir, no perímetro dos seus *poderes de gestão* da sociedade, se esta se faz valer ou não de sistemas de IA que auxiliem/recomendem ou produzam previsões (arts. 192.º, 259.º, 405.º, 474.º e 478.º do CSC). Trata-se de uma decisão de gestão de *natureza discricionária* que versa sobre a "organização dos meios produtivos (…) [e] o sistema informacional inter-orgânico e intra-empresarial"[11].

Através de cláusula(s) estatutária(s), os sócios podem estipular orientações sobre a IA que, designadamente, proíbam a introdução de determinados agentes de IA (art. 6.º, n.º 4, do CSC). Estipulações que, não limitam a capacidade de gozo da sociedade[12], mas "constituem os órgãos da sociedade no dever de não (…) praticarem esses atos" (art. 6.º, n.º 4, do CSC). O desrespeito desta estipulação poderá determinar a responsabilidade civil dos administradores perante a sociedade (art. 72.º do CSC), eventualmente a destituição daqueles, mas, em regra, não obsta à vinculação da sociedade por quotas, anónima e em comandita por ações que, devidamente representada pelo órgão de

---

[10] Cfr. OECD, *Algorithms and Collusion: Competition Policy in the Digital Age*, 2017 www.oecd.org/competition/algorithms-collusion-competition-policy-in-the-digital-age.htm [consulta em 14/4/2022]; Bundeskartellamt/Autorité de la concurrence , *Algorithms and Competition*, November, 2019, disponível em https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2019/06_11_2019_Algorithms_and_Competition.html [consulta em 12/3/2022]; Competition & Markets Authority, *Algorithms: How they can reduce competition and harm consumers*, 19 de janeiro de 2021, disponível em https://www.gov.uk/government/publications/algorithms-how-they-canreduce-competition-and-harm-consumers/algorithms-how-they-can-reduce-competition-and-harmconsumers.

[11] J. M. Coutinho de Abreu, *Governação das sociedades comerciais*, 2.ª ed., Coimbra: Almedina, 2010, 40 (a interpolação não consta do texto original). Por conseguinte, a esta matéria não se aplica a *competência legal residual dos sócios*, prevista no art. 373.º, n.º 2, do CSC, porquanto compete ao conselho de administração gerir as atividades da sociedade (art. 405.º do CSC).

[12] Alexandre de Soveral Martins, «Artigo 6.º - Capacidade», in J. M. Coutinho de Abreu, coord., *Código das Sociedades Comerciais em comentário*, vol. I, 2.ª ed., Coimbra: Almedina, 2017, 133 (117-136). Para a proposta de deslocação dogmática da capacidade jurídica para o âmbito do art. 64.º do CSC, v. Diogo Costa Gonçalves, «Artigo 6.º- Capacidade», in A. Menezes Cordeiro, coord., *Código das Sociedades Comerciais anotado*, 4.ª ed., Coimbra: Almedina, 2022, 117 e ss. (117-127).

administração, adquire tais ferramentas[13]. Em regra, tal aquisição é *eficaz* perante a sociedade por quotas, anónima e em comandita por ações (arts. 260.º, n.os 1, 2 e 3, 409.º, n.os 1, 2 e 3, do CSC)[14].

A decisão de o órgão de administração incorporar ou não IA nas sociedades e, em particular, na gestão das sociedades tem, obrigatoriamente, de ser uma *decisão informada*, tal como exigem os deveres de cuidado contemplados no art. 64.º, n.º 1, *a*), do CSC[15]. E a informação há de abranger não só os benefícios, mas também os riscos que tais tecnologias, por exemplo, a aprendizagem automática não supervisionada, acarretam e os custos inerentes (não só de aquisição, como também de cumprimento[16]). Este é um caso em que é de prever que os administradores necessitem de "auxílio de especialistas internos ou externos"[17] independentes que estão habilitados a produzir informação de qualidade e fiável. Pode ser o caso em que não seja suficiente a informação providenciada pelo produtor ou pela entidade que comercializa a tecnologia.

O cumprimento da lei determinará a recusa de formas de IA inaceitáveis porque atentatórias de "valores da União, como a dignidade do ser humano, a liberdade, a igualdade, a democracia e o Estado de direito, bem como direitos fundamentais da União, incluindo o direito à não discriminação, à proteção de dados pessoais e à privacidade, e os direitos das crianças"[18]. Este será um dos casos em que "os deveres de controlo e vigilância organizativo-funcional podem tornar necessária

---

[13] Admitindo que se trata de sistemas de IA adquiridos não proibidos por lei.

[14] Neste sentido, Alexandre de Soveral Martins, *Os poderes de representação dos administradores de sociedades anónimas*, Coimbra: Coimbra Editora, 1998, 335 (nt. 629) e 337 (n. 633); J. M. Coutinho de Abreu, *Curso de direito comercial*, vol. II. *Das sociedades*, 7.ª ed., Coimbra: Almedina, 2021, 191 e ss.

[15] Alexandre de Soveral Martins, *Administração de sociedades anónimas e responsabilidade dos administradores*, Coimbra: Almedina, 2020, 246, admite a possibilidade de utilização de algoritmos para a análise de *big data*, cujos *outputs* apoiem o processo de decisão.

[16] Para a estimativa de custos de cumprimento, v. Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial), 11.

[17] Pedro Pais de Vasconcelos, «*Business judgment rule*, deveres de cuidado e de lealdade, ilicitude e culpa e o artigo 64.º do Código das Sociedades Comerciais», *Direito das Sociedades em Revista,* 2009, 2, 63.

[18] Cfr. Considerando 15 e Título II da Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial).

a criação de um sistema de *compliance*"[19] que impeça, por exemplo, a "utilização indevida razoavelmente previsível", ou seja, "a utilização de um sistema de IA de uma forma não conforme com a sua finalidade prevista, mas que pode resultar de comportamentos humanos ou de interações com outros sistemas razoavelmente previsíveis"[20].

As vigentes regras de *hard law*[21] relativas à composição e organização do órgão de administração e de representação não lidam especificamente com a emergência de *Corp Tech* e os seus riscos específicos. Em particular, as normas vigentes (sejam elas de *soft law* ou de *hard law*) não garantem que o órgão de administração ou o órgão de fiscalização seja integrado por membros (humanos) especialmente capacitados em IA. Em particular, não está legalmente consagrada a exigência de "tech committees" vocacionados, por exemplo, para a vigilância das negociações com programadores, a revisão das configurações principais dos algoritmos e, possivelmente, a avaliação sobre a remuneração de programadores que trabalham na organização ou, ainda, a supervisão humana de sistemas de IA[22].

O conselho de administração, a não ser que o contrato de sociedade o proíba, pode *encarregar especialmente* algum ou alguns administradores de se ocuparem de matérias relacionadas com a supervisão de IA (art. 407.º, n.º 1, do CSC). O regimento do órgão de administração, enquanto manifestação do poder de auto-organização, pode prever a criação de *comissões atípicas* dedicadas a matérias especificamente atinentes a IA[23].

Pelo seu lado, os sócios poderão estipular cláusulas estatuárias criadoras de *órgãos estatutários*, por exemplo, *comités tecnológicos*. Além

---

[19] Alexandre de Soveral Martins, *Administração de sociedades anónimas e responsabilidade dos administradores*, 224.

[20] Cfr. art. 3.º, 13), da Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial).

[21] Na versão de 2023, o Capítulo VII, relativo ao controlo interno do *Código de Governo das Sociedades* elaborado pelo Instituto Português de *Corporate Governance* inclui a recomendação VII.9., nos termos da qual "a sociedade informa, no relatório de governo, sobre os termos em que mecanismos de inteligência artificial hajam sido utilizados como instrumento de tomada de decisões pelos órgãos sociais".

[22] Luca Enriques / Dirk A. Enriques, *Corporate Technologies and the Tech Nirvana Fallacy*, 46.

[23] Cfr. José Engrácia Antunes, «O regime do órgão de administração», *Direito das Sociedades em Revista* 1/2 (2009) 88 (81-95).

disso, os sócios, através de "cláusula autónoma ou de disposição complementar constante dos estatutos sociais"[24], poderão determinar que a composição do órgão de administração integra membros especialmente habilitados em IA. Considerem-se, ainda, cláusulas estatutárias que fazem depender a introdução de determinadas ferramentas de IA (por exemplo, as classificadas de *risco elevado*) de prévia autorização dos sócios (art. 6.º, n.º 4, do CSC)[25].

Por fim, as sociedades por quotas e anónimas (onde são proibidas as entradas em indústria) podem beneficiar de serviços de sócios especialmente capacitados em IA, prestados no contexto de estipulações estatutárias relativas a *obrigações acessórias*, que tanto podem ser realizadas a título oneroso como gratuito, em regime de trabalho autónomo ou subordinado[26].

3. Alicia T. em Portugal?

Em 2016, a Tieto, uma empresa finlandesa de *software*[27], foi pioneira na nomeação de IA, a *Alicia T.*, para a sua equipa de gestão, também com capacidade para votar. Nesta ocasião, Ari Järvelä, chefe da Tieto Data-Driven Businesses, manifestou a confiança de que a Alicia T. poderia ajudar "a encontrar informação e a tomar decisões baseadas em dados que os humanos não pensam necessariamente – e assim talvez criar algo ainda imprevisto"[28]. No entanto, não há informação

---

[24] Ricardo Costa, «Artigo 391.º - Designação», in J. M. Coutinho de Abreu, coord., *Código das Sociedades Comerciais em comentário*, vol. VI, 2.ª ed., Coimbra: Almedina, 2019, 230 (226-256).

[25] No sentido de que, nas estruturas de administração e de fiscalização previstas no art. 278.º, n.º 1, als. a) e b), do CSC, são lícitas cláusulas relativas ao dever de o conselho de administração obter prévia deliberação dos sócios para a prática de determinadas categorias de atos de gestão, J. M. Coutinho de Abreu, *Governação de sociedades*, 55.

[26] Neste sentido, v. Alexandre da Mota Pinto, «Artigo 209.º - Obrigação de prestações acessórias», in J. M. Coutinho de Abreu, coord., *Código das Sociedades Comerciais em comentário*, vol. III, 3.ª ed., Coimbra: Almedina, 2023, 278.

[27] https://www.tietoevry.com/

[28] Disponível em https://www.bloomberg.com/press-releases/2016-10-17/tieto-the-first-nordic-company-to-appoint-artificial-intelligence-to-the-leadership-team-of-the-new-data-driven-businesses-unit (consultado em 8/3/2022). Tieto, «Tieto: The First Nordic Company to Appoint Artificial Intelligence to the Leadership Team of the New Data-driven Businesses Unit», 2016, disponível em: https://www.tieto.

sobre a atividade de Alicia T. Alguns pensam, até, que este anúncio da Tieto foi mais uma manobra publicitária do que uma efetiva designação de algoritmo para o conselho de administração.

À luz da legislação portuguesa, será *ilícita* e nula a cláusula que suprima órgãos societários consagrados em normas legais imperativas, como são os casos da gerência, ou o órgão de administração e de representação de sociedade anónima e que substitua os "human-populated boards"[29] por sistemas de IA ou que, não suprimindo o órgão de administração e de representação da sociedade, confia a gestão da sociedade a sistemas de IA[30].

As normas sobre a composição e o funcionamento do órgão de administração proíbem também cláusulas estatutárias que deleguem a gestão corrente da sociedade em sistemas de IA (arts. 407.º do CSC)[31].

Uma deliberação dos sócios que numa sociedade por quotas (onde são lícitas as instruções dirigidas pelos sócios aos gerentes, nos termos do art. 259.º do CSC[32]) imponha que a gerência siga decisões tomadas por IA, sem as escrutinar ou vigiar, é certamente nula, por vício de conteúdo (art. 56.º, n.º 1, al. d), CSC), porque o seu conteúdo viola lei imperativa, na medida em que entrega a gestão da sociedade a *entidades não humanas* (arts. 252.º, n.º 1, e 259.º do CSC) ou, ainda, porque viola as regras de competência do órgão de administração. Parece-nos que será nula, porque ilícita, a instrução dirigida pelo

---

com/news/tieto-the-first-nordic-company-to-appoint-artificial-intelligence-to-the-leadership-team-of-the-new [consulta em 15/04/2022].

29 A expressão pertence a Luca Enriques/Dirk A. Zetzsche, *Corporate Technologies and Tech Nirvana Fallacy*, 8.

30 Alexandre de Soveral Martins, «Artigo 405.º - Competência do conselho de administração», in: J. M. Coutinho de Abreu (coord.), *Código das Sociedades Comerciais em comentário*, vol. VI, 2.ª ed., Coimbra: Almedina, 2019, 423 (421-433), que defende que o art. 405.º do CSC "não permite entregar a gestão da sociedade anónima a terceiro".

31 Veja-se, o *Delaware General Corporation Law* (DGCL) §141: "§ 141 Board of directors; powers; number, qualifications, terms and quorum; committees; classes of directors; nonstock corporations; reliance upon books; action without meeting; removal. (a) The business and affairs of every corporation organized under this chapter shall be managed by or under the direction of a board of directors, except as may be otherwise provided in this chapter or in its certificate of incorporation…".

32 Sobre estas, v. Alexandre de Soveral Martins, «Artigo 252.º - Competência da gerência», in J. M. Coutinho de Abreu (coord.), *Código das Sociedades Comerciais em comentário*, vol. IV, 2.ª ed., Coimbra: Almedina, 2017, 144 e ss. (143-149).

órgão de administração da sociedade totalmente dominante ao órgão de administração da sociedade totalmente dominada, no sentido de que este último confie a administração da sociedade a agentes de IA (arts. 481.º, 488.º, 491.º, 503.º, n.º 1, e 504.º, n.º 3, do CSC). Será nula a deliberação de eleição de sistemas de IA como membro(s) do órgão de administração ou cláusula estatutária que os designe como membro do órgão de administração ou de fiscalização, porque estas funções estão reservadas, de modo imperativo, a membros humanos (arts. 191.º, n.os 1 e 3, 252.º, n.º 1, 390.º, n.º 3, 470.º e 56.º, n.º 1, al. d), todos do CSC). E, por conseguinte, constitui-se na esfera jurídica dos administradores (aqui latamente considerados) o dever de não executar tais deliberações nulas[33].

4. Imperfeição da IA e a aplicação da *business judgment rule* a decisões preparadas ou auxiliadas por IA

As consequências danosas de erros causados por decisões preparadas por IA estarão abrangidas pelo art. 72.º, n.º 2, do CSC, influenciado pela *business judgment rule*[34]?

A decisão empresarial de introdução de sistemas de IA é candidata à aplicação da *business judgment rule*, porquanto trata-se de uma decisão empresarial tomada em espaço de discricionariedade, mas, para que tal aconteça, é necessário que os administradores atuem em "termos informados", designadamente, que os administradores se informem sobre a «Finalidade prevista», ou seja, a "utilização à qual o fornecedor destina o sistema de IA, incluindo o contexto específico e as condições de utilização, conforme especificado nas informações facultadas pelo fornecedor nas instruções de utilização, nos materiais e declarações promocionais ou de venda, bem como na documentação

---

[33] Neste sentido, v. Raúl Ventura/Luís Brito Correia, *Responsabilidade civil dos administradores de sociedades anónimas e dos gerentes de sociedades por quotas*, separata do *BMJ*, n.os 192, 193, 194, 195, Lisboa, 1970, 76; J. M. Coutinho de Abreu, *Governação das sociedades comerciais*, 60.

[34] J. M. Coutinho de Abreu/Maria Elisabete Ramos, «Artigo 72.º - Responsabilidade de membros da administração para com a sociedade», in J. M. Coutinho de Abreu, coord., *Código das Sociedades Comerciais em comentário*, vol. I, 2.ª ed., Coimbra: Almedina, 2017, 903 e ss. (892-914).

técnica"[35]. Ou, ainda que se assegurem como funcionam e que riscos comportam os "sistemas de IA utilizados nos domínios do emprego, da gestão de trabalhadores (…), nomeadamente para efeitos de recrutamento e seleção, de tomada de decisões sobre promoções e despedimentos, de repartição de tarefas e de controlo ou avaliação de pessoas no âmbito de relações contratuais de trabalho"[36].

Além disso, o administrador terá de provar que atuou "livre de qualquer interesse pessoal". O conflito de interesses pode existir se, por exemplo, um algoritmo "desenhado" para calcular a remuneração de administradores foi programado para satisfazer os interesses destes, em detrimento do interesse social ou dos sócios[37]. O art. 72.º, n.º 2, do CSC exige *literalmente* que o administrador atue "segundo critérios de racionalidade empresarial". Atualmente, não se pode dizer que seja *irracional* a decisão empresarial de incorporar inteligência artificial não proibida por lei, ainda que seja de elevado risco. Pode acontecer que em tais situações se possa afirmar a violação do dever de tomar decisões razoáveis – e com isso a violação do dever de cuidado (art. 64.º, n.º 1, al. a), do CSC –, *mas não serão irracionais*[38]. E, por isso, o art. 72.º, n.º 2, do CSC poderá impedir a responsabilidade civil dos administradores, ainda que subsistam outras sanções. Tendo em conta os benefícios, os impactos positivos de luta contra a fraude, as vantagens competitivas da inovação e a «corrida» à IA pelas empresas concorrentes, será de concluir que não é, de todo, irracional incorporar IA na empresa.

Os administradores não estarão protegidos pelo art. 72.º, n.º 2,

---

[35] Cfr. art. 3.º, 12), da Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial).

[36] Cfr. o Considerando (36) da Proposta de Regulamento sobre Inteligência Artificial que classifica "Os sistemas de IA utilizados nos domínios do emprego, da gestão de trabalhadores e do acesso ao emprego por conta própria, nomeadamente para efeitos de recrutamento e seleção, de tomada de decisões sobre promoções e despedimentos, de repartição de tarefas e de controlo ou avaliação de pessoas no âmbito de relações contratuais de trabalho" como de "elevado risco", "uma vez que podem ter um impacto significativo nas perspetivas de carreira e na subsistência dessas pessoas".

[37] Luca Enriques/Dirk A. Zetzsche, *Corporate Technologies and the Tech Nirvana Fallacy*, 31, falam em "conflicted coding", para referir a programação adequada (e influenciada pelos) aos interesses da administração da sociedade compradora e não aos interesses da sociedade

[38] J. M. Coutinho de Abreu/Maria Elisabete Ramos, «Artigo 72.º - Responsabilidade de membros da administração para com a sociedade», 906.

do CSC, se as decisões, previsões ou recomendações de agentes de IA violarem a lei porque, por exemplo, são discriminatórias ou violam a privacidade dos consumidores. Por isso, devem ser asseguradas soluções técnicas que, ao longo do ciclo de vida dos agentes de IA, garantam a supervisão humana, a rastreabilidade e a transparência destas ferramentas[39].

Conclusão

A utilização de sistemas de IA nas sociedades é, essencialmente, uma decisão de gestão que deve ser tomada em termos informados. A informação recolhida ou solicitada pelos membros do órgão de administração deve ser adequada ao risco representado pela IA. O que pode implicar a consulta de peritos externos e independentes quer da sociedade quer do fornecedor de IA.

Os deveres de cuidado são suficientemente flexíveis para enquadrarem juridicamente a atuação de administradores que tomam a decisão empresarial de incorporar sistemas de IA na sociedade. As manifestações dos deveres de cuidado determinam, por exemplo, o acompanhamento do desempenho destes sistemas e a preparação da sociedade para a contínua supervisão humana. O que implica que o órgão de administração, no exercício do seu poder de auto-organização, se muna de competências nesta área, como sejam, por exemplo, comissões atípicas para matérias tecnológicas ou encarregar especialmente administrador(es) de questões tecnológicas.

Os deveres de cuidado impõem que os administradores tomem decisões razoáveis, também em matéria de IA. Já o critério que irá avaliar se foi ou não cumprido o dever de tomar decisões razoáveis é o da *irracionalidade* (art. 72º, n.º 2, do CSC). E não se poderá sustentar que, em pleno século XXI, é irracional incorporar IA no funcionamento das sociedades. Já poderá ser ilegal adotar sistemas de IA que

---

[39] O art. 12.º do da Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial), intitulado "manutenção de registos", determina que "Os sistemas de IA de risco elevado devem ser concebidos e desenvolvidos com capacidades que permitam o registo automático de eventos («registos») enquanto o sistema de IA de risco elevado estiver em funcionamento".

sejam proibidos, como se antecipa na Proposta de Regulamento de IA, ou que notoriamente violam a lei. O que mostra que a IA digna de confiança não pode prescindir de uma regulação pública de natureza imperativa.

Referências bibliográficas

Abreu, J. M. Coutinho de, *Governação das sociedades comerciais*, 2.ª ed., Coimbra: Almedina, 2010.

Abreu, J. M. Coutinho de, «Deveres de cuidado e de lealdade dos administradores e interesse social», *Reformas do Código das Sociedades*, Coimbra: Almedina, 2007, 15-47.

Abreu, J. M. Coutinho de *Curso de direito comercial*, vol. II. *Das sociedades*, 7.ª ed., Coimbra: Almedina, 2021.

Abreu, J. M. Coutinho de / Ramos, Maria Elisabete, «Artigo 72.º - Responsabilidade de membros da administração para com a sociedade», in J. M. Coutinho de Abreu, coord., *Código das Sociedades Comerciais em comentário*, vol. I, 2.ª ed., Coimbra: Almedina, 2017, 892-914.

Antunes, José Engrácia, «O regime do órgão de administração», *Direito das Sociedades em Revista* 1/2 (2009) 81-95.

Costa, Ricardo, «Artigo 391.º - Designação», in J. M. Coutinho de Abreu, coord., *Código das Sociedades Comerciais em comentário*, vol. VI, 2.ª ed., Coimbra: Almedina, 2019, 230 (226-256).

Enriques, Luca / Zetzsche, Dirk A., «*Corporate Technologies and the Tech Nirvana Fallacy*», Working paper n.º 457/2019, March 2020, 4, [em linha] [consulta em 12/3/2022]. Disponível em https://ecgi.global/sites/default/files/working_papers/documents/finalenriqueszetzsche.pdf

Martins, Alexandre de Soveral, «Artigo 6.º - Capacidade», in J. M. Coutinho de Abreu, coord., *Código das Sociedades Comerciais em comentário*, vol. I, 2.ª ed., Coimbra: Almedina, 2017, 117-136.

Martins, Alexandre de Soveral, *Os poderes de representação dos administradores de sociedades anónimas*, Coimbra: Coimbra Editora, 1998.

Martins, Alexandre de Soveral, «Artigo 252.º - Competência da gerência», in J. M. Coutinho de Abreu, coord., *Código das Sociedades Comerciais em comentário*, vol. IV, 2.ª ed., Coimbra: Almedina, 2017, 143-149.

Pinto, Alexandre da Mota, «Artigo 209.º - Obrigação de prestações acessórias», in J. M. Coutinho de Abreu, coord., *Código das Sociedades Comerciais em comentário*, vol. III, 3.ª ed., Coimbra: Almedina, 2023.

Ramos Maria Elisabete, "CorpTech e desafios aos deveres de cuidado dos administradores*", Inteligência artificial e robótica desafios para o direito do século XXI*, Sónia Moreira / Pedro Miguel Freitas, coord., Coimbra: Gestlegal, 2022, 229-250, disponível em https://www.jusgov.uminho.pt/pt-pt/publicacoes/inteligencia-artificial-e-robotica-desafios-para-o-direito-do-seculo-xxi/

Ramos, Maria Elisabete / Azevedo, Ana / Meira, Deolinda / Malta, Mariana Curado, «Cooperatives and the use of Artificial Intelligence: A critical view», *Sustainability* 15 (2023) 329. https://doi.org/10.3390/su15010329, disponível em https://www.mdpi.com/journal/sustainability. Special Issue Co-operating for Change: Roles, Potentials, and Challenges of Cooperatives in the Decade Leading up to the Sustainable Development Goals.

Vasconcelos, Pedro Pais de, «*Business judgment rule*, deveres de cuidado e de lealdade, ilicitude e culpa e o artigo 64.º do Código das Sociedades Comerciais», *Direito das Sociedades em Revista* 2 (2009), 41-79.

3

Supervisão, classificação e certificação dos sistemas de IA na Proposta de Regulamento sobre Inteligência Artificial

(https://doi.org/10.47907/DireitoemMudanca/2023/3)

*José Ricardo Marcondes Ramos**

Resumo: Diante das oportunidades e dos riscos oriundos do uso crescente da Inteligência Artificial, a União Europeia vem desenvolvendo medidas e documentos legais para regular o desenvolvimento e uso deste tipo de tecnologia, tendo como objetivo incentivar suas vantagens e ao mesmo tempo tutelar os riscos inerentes. Como resultado desses esforços, foi apresentada a Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial e cria um quadro normativo básico relativo à governação, à supervisão e à responsabilidade. Para consolidar uma abordagem equilibrada que garanta uma gestão eficiente do risco sem prejudicar a inovação, a Proposta de Regulamento da União Europeia sobre inteligência artificial classifica os sistemas de IA em níveis de risco e correlaciona deveres de comportamento específicos e proporcionais a cada tipo. Além disso, também estabelece um conjunto de regras harmonizadas para o desenvolvimento, a colocação no mercado e a utilização de sistemas de IA na UE, e cria requisitos essenciais e obrigatórios para cada tipo de sistema, com foco na transparência e prestação de informações. Neste contexto, o objetivo deste estudo é analisar a Proposta de Regulamento sobre inteligência artificial, e visa examinar os critérios de classificação dos sistemas de IA, as

* Doutorando em Ciências Jurídico-Criminais da Faculdade de Direito da Universidade de Coimbra; investigador-colaborador do IJ.

relações regulatórias que serão instituídas pela legislação na Europa e os deveres de comportamento previstos para mitigar os riscos inerentes aos sistemas de IA.

Palavras-chave: Inteligência artificial, regulação, proposta de regulamento, certificação algorítmica.

1. Introdução

A Inteligência Artificial (IA) já é uma das tecnologias mais transformadoras do século XXI e vem sendo cada vez mais presente no nosso dia a dia, muitas vezes de forma impercetível, seja indicando o melhor caminho no GPS, fazendo uma gestão otimizada da bateria de dispositivos eletrónicos, identificando emails como spam, ou mesmo recomendando conteúdos online. Nos últimos anos, o uso da inteligência artificial tem também trazido uma série de inovações bastante disruptivas a exemplo dos assistentes virtuais como a Alexa ou a Siri, dos sistemas de criação de imagens a partir de texto desenvolvidos por empresas como Dall-E 2, Crayion e Midjourney, e dos Grandes Modelos de Linguagem que baseiam o funcionamento de ChatBots como o Bing da Microsoft, o Bard do Google e o, ChatGPT da Open AI.

Por suas capacidades de melhorar previsões, otimizar as operações e a afetação de recursos de empresas e instituições e personalizar o fornecimento de serviços, o uso da inteligência artificial já tem mostrado que pode contribuir para diversos resultados benéficos para a sociedade e para a economia. Ocorre, porém, que as mesmas habilidades e capacidades técnicas que vêm sendo decisivos para diversos benefícios sociais e económicos têm também demonstrado diversos riscos que representam ameaças à saúde, à segurança e aos direitos fundamentais das pessoas.

Tendo como foco incentivar e maximizar as vantagens que a inteligência artificial pode trazer para a sociedade e, simultaneamente, tutelar os riscos inerentes que esta família de tecnologias pode trazer, a União Europeia vem desenvolvendo uma série de medidas e documentos legais para regular o desenvolvimento e o uso de sistemas de inteligência artificial em âmbito europeu. Levando em consideração a velocidade recente da evolução desta tecnologia e os possíveis desafios sociais daí decorrentes, a UE está empenhada em alcançar uma

abordagem equilibrada que permita garantir uma gestão eficiente do risco de sistemas de inteligência artificial sem prejudicar a inovação. Como fruto destes esforços, foi apresentada pela Comissão Europeia e pelo Parlamento Europeu a Proposta de Regulamento que estabelece regras harmonizadas em matéria de inteligência artificial.

O presente trabalho tem como objeto de estudo a *Proposta de Regulamento* sobre a inteligência artificial, tendo-se como objetivo analisar os critérios de classificação de sistemas de inteligência artificial, as relações regulatórias que serão criadas por esta legislação em âmbito europeu e, finalmente, os deveres de comportamento obrigatórios que os sistemas de inteligência artificial deverão adotar para tutelar os seus riscos inerentes.

2. Classificação de sistemas de Inteligência Artificial

A *Proposta de Regulamento* insere-se em um contexto amplo de regulação da Inteligência Artificial em âmbito Europeu, iniciada ainda em 2018 com a Comunicação da Comissão Europeia *Inteligência Artificial para a Europa (Estratégia IA)* e aprofundada gradualmente em diversos outros documentos como o *Plano coordenado para o desenvolvimento e utilização da inteligência artificial «Made in Europe»* (publicado inicialmente em 2019 e posteriormente revisado em 2021), as *Orientações Éticas para uma IA de Confiança* do grupo de Peritos de Alto Nível sobre a Inteligência Artificial (2019), a Comunicação *Aumentar a Confiança numa Inteligência Artificial Centrada no Ser Humano* (2019) e o *Livro Branco sobre a Inteligência Artificial: uma abordagem europeia virada para a excelência e a confiança* (2020), entre outros.

Como reflexo e consolidação dos princípios e objetivos delineados nos documentos anteriores, a *Proposta de Regulamento* aborda a necessidade de garantir um arcabouço ético e jurídico apropriado para o desenvolvimento da inteligência artificial[1] através da criação um

[1] Enquanto, por exemplo, a Comunicação *Estratégia IA*, de 2018, prevê a garantia de um quadro ético e jurídico apropriado baseado nos valores da União e em consonância com a Carta dos Direitos Fundamentais da União Europeia como um dos 3 pilares bases da abordagem europeia à inteligência artificial (os outros dois pilares são (i) reforçar a capacidade tecnológica e industrial da UE e a aceitação da IA em toda a economia e (ii) preparar a sociedade para as mudanças socioeconómicas decorrentes da IA); em suas *Orientações Éticas para uma IA de Confiança*, o grupo

quadro normativo básico relativo à governação, à supervisão e à responsabilidade em sistemas de inteligência artificial. Seguindo de perto a abordagem europeia centrada no binómio excelência e confiança e no duplo objetivo de promover da adoção da IA enquanto simultaneamente aborda os seus riscos inerentes, delineada no *Livro Branco* que cria as bases da regulação europeia sobre IA, a *Proposta de Regulamento* apresenta um quadro regulamentar horizontal, equilibrado e proporcionado baseado em quatro objetivos específicos, nomeadamente:

- Garantir que os sistemas de IA colocados no mercado da União e utilizados sejam seguros e respeitem a legislação em vigor em matéria de direitos fundamentais e valores da União
- Garantir a segurança jurídica para facilitar os investimentos e a inovação no domínio da IA
- Melhorar a governação e a aplicação efetiva da legislação em vigor em matéria de direitos fundamentais e dos requisitos de segurança aplicáveis aos sistemas de IA
- Facilitar o desenvolvimento de um mercado único para as aplicações de IA legítimas, seguras e de confiança e evitar a fragmentação do mercado.

Para alcançar estes objetivos de forma eficaz e gerir os riscos inerentes aos sistemas de inteligência artificial, a *Proposta de Regulamento* estabelece regras harmonizadas para o desenvolvimento, a colocação no mercado e a utilização de sistemas de IA na União Europeia, o que é feito através da criação de requisitos essenciais e obrigatórios para determinados tipos de sistemas de IA e além de deveres de comportamento relacionados à transparência e à prestação de informações por parte de fornecedores (art. 16), fabricantes (art. 24), importadores (art. 26) e distribuidores (art. 27) de produtos e serviços que utilizem algoritmos de inteligência artificial. Conforme previsto na *Proposta de*

---

de Peritos de Alto Nível sobre a Inteligência Artificial identifica a necessidade de ser Legal, cumprindo toda a legislação e regulamentação aplicáveis como uma das 3 componentes essenciais para uma IA de confiança (as outras duas componentes são (i) ser Ética, garantindo a observância de princípios e valores éticos, e (ii) ser Sólida, tanto do ponto de vista técnico como do ponto de vista social, uma vez que, mesmo com boas intenções, os sistemas de IA podem causar danos não intencionais). De outro lado, o *Livro Branco* ressalta a importância da existência de um quadro regulamentar claro e adaptado às características específicas da inteligência artificial, principalmente a sua capacidade de aprendizagem automática e a sua autonomia decisória.

Regulamento, toda esta estrutura será supervisionada por mecanismos nacionais e europeus de avaliação de conformidade e acompanhamento de sistemas de inteligência artificial.

Consolidando a abordagem europeia centrada no binómio excelência e confiança descrita no *Livro Branco*, a *Proposta de Regulamento* destaca a importância de o quadro regulamentar aplicado ao domínio da inteligência artificial ser eficaz na gestão dos riscos, mas sem ser demasiadamente prescritivo a ponto de inviabilizar a pesquisa e a inovação e também sem criar um encargo desproporcionado principalmente para pequenas e médias empresas[2]. Como forma de criar uma intervenção jurídica equilibrada e proporcional, a *Proposta de Regulamento* propõe um quadro jurídico que é simultaneamente sólido, centrado em uma *abordagem baseada no risco*, que classifica os sistemas de inteligência artificial a partir dos níveis de risco criados pelos sistemas e prevê uma intervenção jurídica às situações concretas em que existe um motivo de preocupação justificado presente ou razoavelmente antecipado num futuro próximo, e simultaneamente flexível, incluindo mecanismos que permitam a sua adaptação dinâmica à medida que a tecnologia evolui e surgem novas situações preocupantes.

Neste contexto, para além de excluir expressamente a sua aplicação aos sistemas de IA desenvolvidos ou utilizados exclusivamente para fins militares (art. 2.º, n. 3), a *Proposta de Regulamento* distingue os sistemas de inteligência artificial em quatro categorias diferentes de risco, correlacionando deveres de comportamento específicos e proporcionais a cada tipo, nomeadamente: *riscos inaceitáveis* (Título II), que são práticas proibidas em território europeu; *riscos limitados* (Título IV), para os quais existem deveres de informação e transparência para com consumidores; *riscos mínimos* (Título IX), para os quais não

---

[2] Conforme descrito no Livro Branco "por uma questão de princípio, o novo quadro regulamentar para a IA deve ser eficaz para atingir os seus objetivos, mas não excessivamente prescritivo, de forma a não criar um encargo desproporcionado, especialmente para as PME. Para atingir este equilíbrio, a Comissão considera que deve seguir uma abordagem baseada no risco". A *Proposta de Regulamento* também reflete esta busca pelo equilíbrio regulatório, estabelecendo em sua exposição de motivos que "a presente proposta apresenta uma abordagem regulamentar horizontal equilibrada e proporcionada ao domínio da inteligência artificial, que se limita aos requisitos mínimos necessários para dar resposta aos riscos e aos problemas associados à IA, sem restringir ou prejudicar indevidamente a evolução tecnológica ou aumentar desproporcionalmente o custo de colocação no mercado das soluções de IA".

existem deveres e obrigações específicos; e, finalmente, *riscos elevados* (Título III), foco principal da *Proposta de Regulamento* para os quais, por ameaçarem direitos fundamentais e a segurança, são previstas uma série de requisitos relacionados à qualidade dos dados, à documentação e à rastreabilidade, à transparência, à supervisão humana, à exatidão e à solidez.

3. Sistemas de inteligência artificial com riscos inaceitáveis

O primeiro tipo de risco identificado pela *Proposta de Regulamento* são os chamados *riscos inaceitáveis* (Título II, art. 5.º), que englobam práticas que a Comissão entende que devem ser proibidas por serem particularmente prejudiciais à população por violarem os valores da União Europeia como a dignidade humana, a liberdade, a igualdade, a democracia e o Estado de direito, bem como direitos fundamentais da União, incluindo o direito à não discriminação, à proteção de dados pessoais e à privacidade, e os direitos das crianças. Neste ponto, a *Proposta de Regulamento* tem como foco quatro tipos de prática, das quais duas relacionadas à manipulação da população com o potencial para causar danos físicos ou psicológicos à pessoa manipulada ou a terceiros, uma relacionada à avaliação e à classificação de pessoas para uso geral por parte das autoridades públicas e, por fim, uma prática relacionada à identificação biométrica à distância em «tempo real» em espaços acessíveis ao público para efeitos de manutenção da ordem pública.

Quanto ao uso de sistemas de inteligência artificial para práticas manipuladoras, exploratórias e de controlo social, por entender que algoritmos concebidos para distorcer o comportamento humano desrespeitam os valores da União, a *Proposta de Regulamento* proíbe a colocação no mercado ou em serviço e a utilização de sistemas de IA (i) baseados em técnicas subliminares que contornem a consciência das pessoas para distorcer substancialmente o seu comportamento (art. 5.º, n. 1, al. a)), bem como algoritmos (ii) que explorem vulnerabilidades de grupos específicos de pessoas associadas à sua idade ou deficiência física ou mental, a fim de distorcer substancialmente o seu comportamento (art. 5.º, n. 1, al. b)).

Em segundo lugar, a *Proposta de Regulamento* proíbe o uso, por parte de autoridades públicas, de sistemas de classificação social que avaliam ou classificam a credibilidade de pessoas singulares com base

no seu comportamento social em diversos contextos ou com base em características de personalidade ou pessoais, conhecidas ou previsíveis (art. 5.º, n. 1, al. c)). Esta proibição decorre do entendimento da Comissão de que este tipo de sistema pode criar resultados discriminatórios, levar a tratamentos prejudiciais, desfavoráveis, injustificados ou desproporcionados de pessoas singulares ou grupos sociais ou mesmo levar à exclusão de grupos inteiros de pessoas, principalmente quando a classificação social obtida por meio desses sistemas for aplicada em contextos sociais não relacionados com o contexto nos quais os dados foram originalmente gerados ou recolhidos (Considerando 17).

A última prática de inteligência artificial proibida pela *Proposta de Regulamento* relaciona-se ao uso de sistemas de identificação biométrica de pessoas singulares à distância que, na definição adotada no artigo 3.º, n. 36, pode ser compreendido como "um sistema de IA concebido para identificar pessoas singulares à distância por meio da comparação dos dados biométricos de uma pessoa com os dados biométricos contidos numa base de dados de referência, sem que o utilizador do sistema de IA saiba antecipadamente se a pessoa em causa estará presente e pode ser identificada". Diferentemente das proibições anteriores, a vedação trazida no artigo 5.º, n. 1, al. d) da *Proposta de Regulamento*, não está relacionada ao sistema de inteligência artificial em si mesmo, mas à forma como este tipo de sistema pode vir a ser utilizado, motivo pelo qual a proibição está relacionada à verificação simultânea de três elementares normativas principais: que a identificação biométrica à distância seja feita «em tempo real», realizada em espaços acessíveis ao público e com o propósito específico de realizar a manutenção da ordem pública.

Com relação ao primeiro elemento, a *Proposta de Regulamento* diferencia a identificação biométrica à distância «em tempo real» e «em diferido»[3] descrevendo a primeira forma como "a utilização «ao vivo» ou «quase ao vivo» de materiais, como vídeos, criados por uma câmara ou outro dispositivo com uma funcionalidade semelhante"

---

[3] Em contraste, o artigo 3.º, n. 37 define os sistemas de identificação biométrica à distância em diferido como "um sistema de identificação biométrica à distância que não seja um sistema de identificação biométrica à distância em «tempo real»". De outro lado, porém, o considerando 8 esclarece que "no caso dos sistemas «em diferido», os dados biométricos já foram recolhidos e a comparação e a identificação ocorrem apenas após um atraso significativo. Estes sistemas utilizam materiais, tais como imagens ou vídeos, criados por câmaras de televisão em circuito fechado ou dispositivos privados antes de o sistema ser utilizado relativamente às pessoas singulares em causa".

(Considerando n.º 8), ou como a recolha de dados biométricos, a comparação com uma base de dados de referência e a identificação de pessoas singulares sem atraso significativo, de forma imediata ou quase imediata, ou em todo o caso, sem um atraso significativo (art. 3.º, n. 36). Diante da importância desta elementar normativa, o texto legal ainda inclui a previsão expressa de que "[p]ara evitar que as regras sejam contornadas, tal inclui não apenas a identificação instantânea, mas também a identificação com ligeiro atraso".

A delimitação da proibição aos espaços acessíveis ao público – definido pelo artigo 3.º, n. 39) como "qualquer espaço físico aberto ao público, independentemente da eventual aplicação de condições de acesso específicas" – está relacionada não apenas à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais mas, principalmente, à gestão dos riscos inerentes à utilização de sistemas de identificação biométrica de pessoas singulares. Isto porque a Comissão Europeia entende que a utilização deste tipo de sistema "é considerada particularmente intrusiva para os direitos e as liberdades das pessoas em causa, visto que pode afetar a vida privada de uma grande parte da população, dar origem a uma sensação de vigilância constante e dissuadir indiretamente o exercício da liberdade de reunião e de outros direitos fundamentais" (Considerando 19). Além disso, a Comissão também ressalta que existem riscos acrescidos para os direitos e as liberdades das pessoas decorrentes do impacto imediato e das oportunidades limitadas para a realização de controlos adicionais ou correções da utilização desses sistemas que funcionam «em tempo real».

Finalmente, a última elementar normativa que interessa a esta proibição está relacionada ao contexto em que os sistemas de identificação biométrica à distância são aplicados, isto é, para efeitos de manutenção da ordem pública o que é compreendido como "atividades realizadas por autoridades policiais para efeitos de prevenção, investigação, deteção ou repressão de infrações penais ou execução de sanções penais, incluindo a proteção contra ameaças à segurança pública e a prevenção das mesmas" (art. 3.º, n. 41). Esta elementar é relevante uma vez que é precisamente no contexto da manutenção da ordem pública que a *Proposta de Regulamento* descreve três situações de exceção em que a utilização de sistemas de identificação biométrica de pessoas singulares à distância em tempo real é permitida, nomeadamente, a procura e investigação seletiva de potenciais vítimas específicas de crimes,

principalmente crianças desaparecidas; a prevenção de uma ameaça específica, substancial e iminente à vida ou à segurança física de pessoas singulares ou de um ataque terrorista; e, finalmente, a deteção, localização, identificação ou instauração de ações penais relativamente a infratores ou suspeitos de infrações penais a que se refere a Decisão-Quadro 2002/584/JAI do Conselho[4], desde que tal infração seja punível no Estado-Membro em causa com pena ou medida de segurança privativas de liberdade de duração máxima não inferior a três anos e tal como definidas pela legislação desse Estado-Membro.

Muito embora estas três exceções sejam permitidas, uma vez que a Comissão entende que são situações em que a utilização de sistemas de identificação biométrica de pessoas singulares à distância em tempo real é estritamente necessária por motivos de interesse público importante e cuja importância prevalece sobre os seus riscos inerentes, a *Proposta de Regulamento* condiciona esta utilização à autorização expressa e específica de uma autoridade judiciária competente ou de uma autoridade administrativa independente. Ainda assim, apesar desta exigência de autorização jurisdicional prévia, a *Proposta de Regulamento* também prevê a possibilidade da apresentação de um pedido durante o uso do sistema ou logo após, desde que haja uma situação de emergência devidamente justificada, "ou seja, quando a necessidade de utilizar os sistemas em causa torna efetiva e objetivamente impossível obter uma autorização antes de iniciar essa utilização" (Considerando

---

[4] As infrações em questão são as seguintes: participação numa organização criminosa; terrorismo; tráfico de seres humanos; exploração sexual de crianças e pedopornografia; tráfico ilícito de estupefacientes e de substâncias psicotrópicas; tráfico ilícito de armas, munições e explosivos; corrupção; fraude, incluindo a fraude lesiva dos interesses financeiros das Comunidades Europeias na aceção da convenção de 26 de julho de 1995, relativa à proteção dos interesses financeiros das Comunidades Europeias; branqueamento dos produtos do crime; falsificação de moeda, incluindo a contrafação do euro; cibercriminalidade; crimes contra o ambiente, incluindo o tráfico ilícito de espécies animais ameaçadas e de espécies e essências vegetais ameaçadas; auxílio à entrada e à permanência irregulares; homicídio voluntário, ofensas corporais graves; tráfico ilícito de órgãos e de tecidos humanos; rapto, sequestro e tomada de reféns; racismo e xenofobia; roubo organizado ou à mão armada; tráfico de bens culturais incluindo antiguidades e obras de arte; burla; extorsão de proteção e extorsão; contrafação e piratagem de produtos; falsificação de documentos administrativos e respetivo tráfico; falsificação de meios de pagamento; tráfico ilícito de substâncias hormonais e outros fatores de crescimento; tráfico ilícito de materiais nucleares e radioativos; tráfico de veículos roubados; violação; fogo-posto; crimes abrangidos pela jurisdição do Tribunal Penal Internacional; desvio de avião ou navio; sabotagem.

21). Ainda assim, segundo previsão do texto legal, o uso dos sistemas de identificação biométrica em situações de emergência deve limitar-se ao mínimo absolutamente necessário e a autoridade policial deve obter uma autorização o quanto antes, apresentando as razões para não ter efetuado o pedido mais cedo.

4. Sistemas de inteligência artificial de risco limitado e de risco mínimo

Entre os sistemas de inteligência artificial cuja utilização é permitida em âmbito Europeu, a *Proposta de Regulamento* identifica em seu artigo 52.º (Título IV) três tipos de sistemas de IA que geram *riscos limitados* para os quais são previstas regras de transparência harmonizadas que criam deveres de prestação de informações para seus usuários, de forma a garantir que as pessoas possam tomar decisões informadas ou distanciar-se de determinadas situações – o que inclui a obrigação de que essas informações e notificações devem ser fornecidas em formatos acessíveis a pessoas com deficiência. Em específico, estas obrigações de transparência são aplicáveis a (i) sistemas autónomos que interagem com pessoas singulares (art. 52.º, n. 1), (ii) sistemas de reconhecimento de emoções ou de categorização biométrica (art. 52.º, n. 2), e (iii) sistemas de geração ou manipulação de conteúdo (art. 52.º, n. 3).

Inicialmente, por entender que sistemas autónomos que interagem com pessoas singulares e sistemas de criação e manipulação de conteúdo "podem representar riscos específicos de usurpação de identidade ou fraude" (Considerando 70), a *Proposta de Regulamento* cria obrigações de transparência específicas. Assim, em primeiro lugar, fornecedores de sistemas de IA destinados a interagir com seres humanos têm o dever de garantir que este tipo de algoritmo seja concebido e desenvolvido de maneira a que as pessoas sejam informadas de que estão a interagir com um sistema de IA e não com outro ser humano, salvo quando as circunstâncias e o contexto de utilização do algoritmo revelem esta interação automatizada de forma óbvia[5] (art. 52.º, n. 1).

---

[5] Cumpre salientar que a *Proposta de Regulamento* prevê uma segunda exceção estabelecendo que "esta obrigação não se aplica a sistemas de IA legalmente autorizados para detetar, prevenir, investigar e reprimir infrações penais, salvo se esses sistemas estiverem disponíveis ao público para denunciar uma infração penal".

Em segundo lugar, utilizadores que recorram a sistemas de inteligência artificial para gerar ou manipular conteúdos de imagem, áudio ou vídeo que sejam consideravelmente semelhantes a conteúdos autênticos – ou seja, que sejam consideravelmente semelhantes a pessoas, locais ou acontecimentos reais e que, falsamente, pareçam ser autênticos a outrem, as chamadas «falsificações profundas» – têm o dever de informar aos usuários que o conteúdo em questão foi gerado ou manipulado artificialmente (art. 52.º, n. 3). O texto legislativo prevê como exceção a estas obrigações de transparência a utilização destes sistemas para fins legítimos, descrevendo duas situações específicas: a sua utilização no contexto da manutenção da ordem pública e em casos de liberdade de expressão.

Finalmente, os deveres de transparência e informação são também aplicáveis a utilizadores de sistemas de inteligência artificial de reconhecimento de emoções e de categorização biométrica de pessoas singulares (art. 52.º, n. 2). Assim, ainda como forma de garantir que as pessoas possam tomar decisões informadas ou mesmo distanciar-se de determinadas situações, existe uma obrigação de informação aos usuários quando as suas emoções ou características são reconhecidas por meios automatizados ou quando estão sujeitas à avaliação e associação com categorias (sociais) com base em dados biométricos. Aqui, a única exceção em que esta obrigação não se aplica relaciona-se à utilização de sistemas de IA usados para categorização biométrica que sejam legalmente autorizados para detetar, prevenir e investigar infrações penais.

Relativamente aos deveres de transparência dos sistemas de inteligência artificial de risco limitado, é interessante notar que enquanto os deveres de comportamento relacionados aos sistemas de reconhecimento de emoções e categorização biométrica e aos sistemas de manipulação de conteúdo aplicam-se aos seus *utilizadores* – que na aceção do artigo 3.º, n. 4, compreende "uma pessoa singular ou coletiva, autoridade pública, agência ou outro organismo que utilize, sob a sua autoridade, um sistema de IA, salvo se o sistema de IA for utilizado no âmbito de uma atividade pessoal de caráter não profissional" – as obrigações de informação relacionadas aos sistemas automáticos de interação (ou *chatbots*) aplicam-se aos *fornecedores*, compreendidos como "uma pessoa singular ou coletiva, autoridade pública, agência ou outro organismo que desenvolva um sistema de IA ou que tenha um sistema de IA desenvolvido com vista à sua colocação no mercado ou

colocação em serviço sob o seu próprio nome ou marca, a titulo oneroso ou gratuito" (art. 3.º, n. 2).

Relativamente aos sistemas de inteligência artificial de risco mínimo, a *Proposta de Regulamento* não menciona este tipo de sistema diretamente, motivo pelo qual eles são identificados de forma subsidiária por não estarem em nenhuma das classificações anteriores. Muito embora para este tipo de sistema não existam obrigações legais aplicáveis ou deveres de comportamento específicos, por força do Título IX da *Proposta de Regulamento* a Comissão Europeia encoraja e facilita a adoção de códigos de conduta voluntários por todos os sistemas de inteligência artificial utilizados na União Europeia, independentemente dos deveres de comportamento específicos correlacionados a cada tipo de risco.

5. Sistemas de inteligência artificial de risco elevado

Para além das obrigações de transparência e informação relacionadas a sistemas de risco limitado, a *Proposta de Regulamento* cria também um conjunto de requisitos horizontais obrigatórios que garantam uma IA de confiança, procedimentos de avaliação de conformidade antes da sua colocação no mercado e obrigações previsíveis, proporcionadas e claras para garantir a segurança e o respeito da legislação em vigor. Ainda como reflexo da abordagem baseada no risco e seguindo o objetivo de criar uma intervenção jurídica equilibrada e proporcional, a *Proposta de Regulamento* descreve estes deveres de comportamento específicos como "estritamente necessários para atenuar os riscos" colocados pela inteligência artificial nos domínios da saúde, segurança e direitos fundamentais, restringindo esta intervenção às situações concretas em que existe um motivo de preocupação justificado presente ou razoavelmente antecipado num futuro próximo, naquilo que identifica como *sistemas de inteligência artificial de risco elevado* (Capítulo III)[6].

---

[6] Conforme descrito no Considerando 28: "a dimensão dos impactos adversos causados pelo sistema de IA nos direitos fundamentais protegidos pela Carta é particularmente importante quando se classifica um sistema de IA como sendo de risco elevado. Esses direitos incluem o direito à dignidade do ser humano, o respeito da vida privada e familiar, a proteção de dados pessoais, a liberdade de expressão e de informação, a liberdade de reunião e de associação, a não discriminação, a defesa dos consumidores, os direitos dos trabalhadores, os direitos das pessoas com deficiência, o

Reconhecendo que o condicionamento da introdução de bens e serviços no mercado europeu ao cumprimento de deveres específicos implica restrições à liberdade de empresa e à liberdade das artes e das ciências, em seu artigo 6.º (Título III, Capítulo 1) a *Proposta de Regulamento* delimita duas regras claras para a classificação de um sistema de inteligência artificial como sendo de risco elevado, especificamente: (i) caso o sistema de IA destine-se a ser utilizado como componente de segurança de um produto, ou seja ele mesmo um produto que já é objeto de um procedimento de avaliação da conformidade por força da legislação da União Europeia; e (ii) caso seja um sistema de IA autónomo que, em função da sua finalidade prevista, represente um risco elevado de danos para a saúde e a segurança ou de prejuízo para os direitos fundamentais das pessoas, tendo em conta a gravidade dos possíveis danos e a probabilidade dessa ocorrência, e se forem utilizados num conjunto de domínios especificamente predefinidos no regulamento.

Primeiramente, com relação aos sistemas de IA que constituem componentes de segurança ou são eles mesmos produtos já objeto de avaliação e certificação, a aplicação de deveres de comportamento e a obrigatoriedade de certificação para a segurança dos consumidores têm como foco garantir a integração da regulação da inteligência artificial à legislação de segurança setorial em vigor, assegurando a coerência legislativa, evitando duplicações e minimizando os encargos adicionais. Assim sendo, tendo em consideração a lista de produtos já sujeitos à certificação por força da legislação europeia setorial, descrita no Anexo II da *Proposta de Regulamento*, tem-se que são classificados como sistemas de IA de risco elevado os algoritmos utilizados como sistemas de segurança dos seguintes produtos: máquinas, brinquedos, ascensores, aparelhos e sistemas de proteção destinados a ser utilizados em atmosferas potencialmente explosivas, equipamentos de rádio, equipamentos sob pressão, equipamentos de embarcações de recreio, instalações por cabo, aparelhos a gás, dispositivos médicos e dispositivos médicos para diagnóstico *in vitro*.

Na medida em que a *Proposta de Regulamento* esclarece que a classificação deste tipo de sistema como de risco elevado tem como objetivo "prevenir e atenuar devidamente os riscos de segurança que possam ser

---

direito à ação e a um tribunal imparcial, a presunção de inocência e o direito de defesa e o direito a uma boa administração".

criados por um produto devido aos seus componentes digitais, incluindo sistemas de IA" (Considerando 28), como consequência da integração da *Proposta de Regulamento* ao arcabouço legislativo europeu de proteção aos consumidores, tem-se que a verificação do cumprimento dos requisitos aplicáveis aos sistemas de IA será realizado no âmbito do chamado *novo quadro legislativo* (NQL), com a avaliação e a certificação dos requisitos obrigatórios descritos na *Proposta de Regulamento* em conjunto com os mecanismos de conformidade e execução *ex ante* e *ex post* aplicáveis aos produtos acima identificados[7].

Relativamente ao segundo tipo de sistemas de inteligência artificial classificado como de risco elevado, a *Proposta de Regulamento* esclarece que esta classificação é apropriada se, em função da finalidade prevista, os algoritmos "representarem um risco elevado de danos para a saúde e a segurança ou de prejuízo para os direitos fundamentais das pessoas, tendo em conta a gravidade dos possíveis danos e a probabilidade dessa ocorrência, e se forem utilizados num conjunto de domínios especificamente predefinidos no regulamento" (Considerando 32). Desta forma, em seu Anexo III, a *Proposta de Regulamento* traz uma lista exaustiva com um número limitado de sistemas de IA cujos riscos já se materializaram ou são suscetíveis de se materializar num futuro próximo, incluídos em um dos seguintes domínios[8]: identificação biométrica e caracterização de pessoas singulares à distância em tempo real ou em diferido, nas exceções autorizadas; gestão e funcionamento de infraestruturas críticas, como trânsito rodoviário e redes de abastecimento de água, gás, aquecimento e eletricidade; educação e formação profissional; acesso a serviços privados e públicos essenciais; manutenção

---

[7] Segundo descreve a Comissão "a principal diferença é que os mecanismos de *ex ante* e *ex post* assegurarão o cumprimento não só dos requisitos estabelecidos pela legislação setorial, mas também dos requisitos estabelecidos pelo presente regulamento".

[8] Por força do artigo 7.º da *Proposta de Regulamento,* a Comissão Europeia mantém poderes para adotar Atos Delegados (art. 73.º) para atualizar a lista do Anexo III, aditando os sistemas de IA considerados de risco elevado, desde que preencham cumulativamente 2 requisitos: que os sistemas de IA destinem-se a ser utilizados em qualquer um dos domínios enumerados no anexo III, pontos 1 a 8 e que representem um risco de danos para a saúde e a segurança ou um risco de impacto adverso nos direitos fundamentais que, em termos de gravidade e probabilidade de ocorrência, é equivalente ou superior ao risco de danos ou impacto adverso representado pelos sistemas de IA de risco elevado já referidos no anexo III.

da ordem pública; gestão de migração, asilo e controlo de fronteiras; e administração da justiça e processos democráticos.

Com relação aos sistemas de IA concebidos para a identificação biométrica à distância de pessoas singulares, a Comissão esclarece que este tipo de sistema deve ser considerado de risco elevado na medida em que pode conduzir a resultados enviesados e ter efeitos discriminatórios, particularmente no que diz respeito à idade, à etnia, ao sexo ou a deficiências das pessoas. Assim, tendo em consideração seus riscos inerentes, os sistemas de identificação biométrica, tanto em sua forma «em tempo real» quanto «em diferido», devem estar sujeitos a requisitos específicos relativos às capacidades de registo e à supervisão humana (Considerando 33). Aliás, o cuidado da Comissão com este tipo de sistema é tal que os sistemas de identificação biométrica de pessoas é o único tipo de sistema de inteligência artificial de risco elevado cuja avaliação de conformidade não poderá ser feita pelo próprio fornecedor, sendo obrigatoriamente necessária a participação de um organismo notificado (Considerando 64).

De outro lado, a utilização de sistemas de IA na gestão e funcionamento de infraestruturas críticas como componentes de segurança no controlo do tráfego rodoviário e das redes de abastecimento de água, gás, aquecimento e eletricidade, é tida como de risco elevado uma vez que a Comissão entende que uma falha ou anomalia nestes sistemas pode pôr em risco a vida e a saúde das pessoas em larga escala e provocar perturbações substanciais das atividades sociais e económicas normais (Considerando 34).

Outro domínio no qual sistemas de inteligência artificial são considerados de risco elevado é o da educação ou formação profissional, designadamente quando utilizados para determinar o acesso ou a afetação de pessoas a instituições de ensino e de formação profissional ou para avaliar testes que as pessoas realizam no âmbito da sua educação ou como pré-condição para a mesma. Aqui, a Comissão destaca no Considerando 35 que o uso de sistemas de inteligência artificial possui um risco inerente uma vez que, caso sejam indevidamente concebidos e utilizados, estes sistemas podem violar o direito à educação e à formação, bem como o direito a não ser alvo de discriminação e de perpetuação de padrões históricos de discriminação, o que possui importância central para a sociedade na medida em que este tipo de violação influencia diretamente o percurso académico e profissional

das pessoas e, por consequência, a sua capacidade de garantir a própria subsistência.

Refletindo e complementando a tutela anterior da capacidade de subsistência dos trabalhadores da União Europeia, também são considerados de risco elevado os sistemas de inteligência artificial aplicados ao domínio do emprego, da gestão de trabalhadores e de acesso ao emprego por conta própria. Na medida em que a utilização de sistemas de IA como assistente decisório pode levar recrutadores, empregadores e gestores de recursos humanos a decisões erradas ou enviesadas que perpetuam padrões históricos de discriminação (por exemplo, contra as mulheres, certos grupos etários, pessoas com deficiência ou pessoas de uma determinada origem racial ou étnica ou orientação sexual), a *Proposta de Regulamento* procura gerir o uso da inteligência artificial em dois âmbitos principais de relações trabalhistas: a seleção e recrutamento de novos trabalhadores, designadamente para divulgação de vagas, aplicações de triagem ou filtragem de currículos, e avaliação de candidatos; e na avaliação de desempenho durante a vigência de relações de trabalho, na tomada de decisões sobre promoções ou despedimentos, sobre a repartição de tarefas e, por fim, no controlo e avaliação do desempenho e do comportamento dos trabalhadores (Considerando 36).

Também sob a perspetiva de evitar a ocorrência de discriminação de pessoas ou grupos, evitar a criação de impactos discriminatórios e impedir a perpetuação de padrões históricos de discriminação em razão da origem étnica ou racial, deficiência, idade ou orientação sexual, ou criar novas formas de impactos discriminatórios, são classificados como sendo de risco elevado os sistemas de inteligência artificial utilizados para gerir o acesso a determinados serviços e prestações essenciais, tanto de cariz privado quanto público, e o usufruto dos mesmos. Segundo destaca a Comissão (Considerando 37), a classificação destes sistemas como sendo de risco elevado decorre do impacto potencial que sistemas discriminatórios podem trazer para as pessoas e da importância dos serviços públicos e privados para que as pessoas participem plenamente na sociedade ou melhorem o seu nível de vida.

No âmbito dos serviços privados, o foco da *Proposta de Regulamento* é o acesso da população ao setor financeiro e à possibilidade das pessoas de terem acesso a recursos financeiros ou a serviços essenciais, como o alojamento, a eletricidade e os serviços de telecomunicações, motivo pelo qual são considerados de risco elevado os sistemas de IA

concebidos para avaliar a classificação de crédito ou a capacidade de endividamento de pessoas singulares[9]. De outro lado, relativamente à utilização de sistemas de IA para gerir o acesso aos serviços públicos, ainda que reconheça a importância de evitar a criação de obstáculos ao desenvolvimento e à utilização de abordagens inovadoras na administração pública, a *Proposta de Regulamento* destaca a importância dos cuidados no desenvolvimento deste tipo de algoritmo, na medida em que as pessoas singulares que se candidatam ou que recebem prestações e serviços de assistência pública dependem dos mesmos e estão numa posição vulnerável face às autoridades competentes motivo pelo qual os sistemas de IA podem ter um impacto significativo na subsistência das pessoas e podem infringir os seus direitos fundamentais, como o direito à proteção social, à não discriminação, à dignidade do ser humano ou à ação.

Por este motivo, são considerados como de risco elevado não apenas os algoritmos concebidos para serem utilizados no envio ou no estabelecimento de prioridades no envio de serviços de resposta a emergências, incluindo bombeiros e assistência medica, mas também os sistemas de IA concebidos para serem utilizados por autoridades públicas ou em nome de autoridades públicas para avaliar a elegibilidade de pessoas singulares quanto a prestações e serviços públicos de assistência, bem como para conceder, reduzir, revogar ou recuperar tais prestações e serviços.

Ainda no âmbito da administração pública, outro tipo de sistema considerado de risco elevado são os algoritmos de inteligência artificial utilizados em ações das autoridades policiais para a manutenção da ordem pública, contexto social particularmente sensível por ser caracterizado por um grau substancial de desequilíbrio de poder e que pode conduzir à vigilância, detenção ou privação da liberdade de uma pessoa singular, bem como ter outros impactos adversos nos direitos fundamentais garantidos pela Carta. Neste ponto, a exatidão, a fiabilidade e a transparência dos sistemas de inteligência artificial possuem

---

[9] A única exceção aplicável são sistemas de avaliação de crédito desenvolvidos e utilizados por fornecedores de pequena dimensão: "Tendo em conta a dimensão bastante limitada do impacto e as alternativas disponíveis no mercado, é apropriado isentar os sistemas de IA utilizados para efeitos de avaliação da capacidade de endividamento e de classificação de crédito que sejam colocados em serviço por fornecedores de pequena dimensão para utilização própria" (Considerando 37).

importância central para evitar impactos adversos na sociedade, reter a confiança do público e assegurar a responsabilidade e vias de recurso eficazes, particularmente porque, primeiro, se não for treinado com dados de alta qualidade, não cumprir os requisitos adequados em termos de exatidão ou solidez ou não tiver sido devidamente concebido e testado antes de ser colocado no mercado ou em serviço, o sistema de IA pode destacar pessoas de uma forma discriminatória ou incorreta e injusta; e, segundo, o exercício de importantes direitos fundamentais processuais, como o direito à ação e a um tribunal imparcial, a presunção de inocência e o direito de defesa, pode ser prejudicado[10] (Considerando 28).

Assim, tendo em conta a natureza das atividades em causa e os riscos associados às mesmas, serão considerados de risco elevado os sistemas de inteligência artificial concebidos para serem utilizados pelas autoridades policiais (i) em avaliações individuais de riscos (tanto para determinar o risco de uma pessoa singular cometer infrações ou voltar a cometer infrações ou quanto para avaliar o risco para potenciais vítimas de infrações penais), (ii) em instrumentos utilizados para detetar o estado emocional de uma pessoa singular (como polígrafos ou instrumentos semelhantes), (iii) para detetar «falsificações profundas», (iv) para avaliar a fiabilidade dos elementos de prova em processos penais, (v) para prever a ocorrência ou a recorrência de uma infração penal real ou potencial com base na definição de perfis de pessoas singulares ou para avaliar os traços de personalidade e as características ou o comportamento criminal passado de pessoas singulares ou grupos, e (vi) para a definição de perfis no decurso da deteção, investigação ou repressão de infrações penais, bem como para o estudo analítico de crimes relativos a pessoas singulares.

O penúltimo domínio de sistema de inteligência artificial aplica-se igualmente à administração pública, especificamente relacionado à utilização na gestão da migração, do asilo e do controlo de fronteiras[11].

---

[10] Segundo prevê a *Proposta de Regulamento, "*os sistemas de IA especificamente concebidos para serem utilizados em processos administrativos por autoridades fiscais e aduaneiras não devem ser considerados sistemas de IA de risco elevado utilizados por autoridades policiais para efeitos de prevenção, deteção, investigação e repressão de infrações penais" (Considerando 38).

[11] Para além de classificar este tipo de sistema de IA como de risco elevado, como forma de assegurar a integração da certificação deste tipo de algoritmo à legislação

Segundo descreve a Comissão (Considerando 39), por tratar-se de um âmbito que afeta pessoas que, via de regra, encontram-se numa posição particularmente vulnerável e que dependem do resultado das ações das autoridades públicas competentes, a exatidão, a natureza não discriminatória e a transparência dos sistemas de IA utilizados nesses contextos são particularmente importantes para garantir o respeito dos direitos fundamentais das pessoas em causa, nomeadamente os seus direitos à livre circulação, à não discriminação, à proteção da vida privada e dos dados pessoais, à proteção internacional e a uma boa administração.

Com isso, são considerados de risco elevado os sistemas de inteligência artificial utilizados neste domínio para detetar o estado emocional de uma pessoa singular (como polígrafos ou instrumentos similares), para avaliar determinados riscos colocados pelas pessoas singulares que entram no território de um Estado-Membro ou pedem um visto ou asilo; para verificar a autenticidade dos documentos apresentados pelas pessoas singulares; para auxiliar as autoridades públicas competentes na análise dos pedidos de asilo, de visto e de autorização de residência e das queixas relacionadas, com o objetivo de estabelecer a elegibilidade das pessoas singulares que requerem determinado estatuto.

Finalmente, o último tipo de sistema de inteligência artificial considerado de risco elevado são aqueles aplicáveis à administração da justiça e aos processos democráticos, ou seja, algoritmos concebidos para auxiliar as autoridades judiciárias na investigação e na interpretação de factos e do direito e na aplicação da lei a um conjunto específico de factos. O foco central da *Proposta de Regulamento* está associado ao potencial de impacto negativo significativo que enviesamentos, erros e opacidade de sistemas de IA podem trazer à democracia, ao Estado de direito e às liberdades individuais, bem como ao direito à ação e a um tribunal imparcial. Tendo em conta os riscos específicos que o texto legislativo busca tutelar, sistemas de IA concebidos para atividades administrativas puramente auxiliares que não afetam a administração efetiva da justiça em casos individuais (a exemplo da anonimização ou a pseudonimização de decisões judiciais, documentos ou dados,

---

setorial em vigor, a *Proposta de Regulamento* ainda determina que os sistemas de IA aplicáveis no domínio da gestão da migração, do asilo e do controlo das fronteiras devem cumprir os requisitos processuais estabelecidos na Diretiva 2013/32/UE do Parlamento Europeu e do Conselho, no Regulamento (CE) n.º 810/2009 do Parlamento Europeu e do Conselho e noutra legislação aplicável.

comunicações entre pessoal, tarefas administrativas ou afetação de recursos) não estão abrangidas por esta classificação.

6. Supervisão e certificação de sistemas de inteligência artificial de risco elevado

Tendo em conta os riscos potenciais à sociedade criados pelos sistemas de inteligência artificial de risco elevado e como forma de tutelar eventuais danos para a saúde, a segurança ou para os direitos fundamentais das pessoas, a *Proposta de Regulamento* cria a obrigatoriedade de que este tipo de sistema cumpra um conjunto de requisitos obrigatórios horizontais para uma IA de confiança, que serão subsequentemente operacionalizados por via de normas técnicas harmonizadas. Como forma de integrar a supervisão dos sistemas de inteligência artificial ao arcabouço legislativo da União Europeia de proteção aos consumidores, de segurança dos produtos e de garantia da livre circulação de produtos e serviços no mercado europeu, a *Proposta de Regulamento* segue o modelo já utilizado no âmbito do *novo quadro legislativo* (NQL), instituído para a avaliação e certificação de determinados produtos, e cria dois tipos de relações regulatórias relacionados, primeiro, à testagem, prestação de informações e documentação *antes* da colocação de um sistema de IA no mercado e, segundo, de controlo, manutenção de registos e prestação de informações sobre incidentes graves ou anomalias no *pós-comercialização*, durante todo o ciclo de vida do algoritmo.

Primeiramente, como forma de assegurar um nível elevado de fiabilidade dos sistemas de IA de risco elevado perante os consumidores europeus, antes de serem colocados no mercado ou em serviço estes sistemas deverão passar por um procedimento de *avaliação de conformidade* (art. 19), que vai verificar o cumprimento de todos os requisitos obrigatórios determinados pelo Capítulo 2, Título III, da *Proposta de Regulamento* (Considerando 62). Conforme aponta a Comissão Europeia, estas "obrigações relativas à testagem *ex ante*, à gestão de riscos e à supervisão humana também facilitarão o respeito de outros direitos fundamentais, graças à minimização do risco de decisões assistidas por IA erradas ou enviesadas em domínios críticos como a educação e a formação, o emprego, serviços essenciais, a manutenção da ordem pública e o sistema judicial".

Aqui, surge um traço característico do sistema europeu: na busca de criar uma intervenção jurídica equilibrada e proporcional e como forma de minimizar os encargos impostos aos desenvolvedores, especialmente no caso de pequenas e médias empresas, esta avaliação de conformidade será uma autoavaliação realizada pelo fornecedor sob a sua própria responsabilidade (Considerando 64), que vai consolidar o cumprimento de todos os requisitos obrigatórios em uma documentação técnica (art. 11). Como já se mencionou, a *Proposta de Regulamento* prevê uma única exceção aplicável aos sistemas de IA concebidos para serem utilizados para a identificação biométrica à distância de pessoas singulares, caso em que os sistemas deverão sempre ser certificados pelos chamados organismos notificados que verificarão o cumprimento de todos os requisitos técnicos obrigatórios (Considerando 65).

Conforme explicitado no artigo 43.º da *Proposta de Regulamento* e detalhado em seu Anexo VI, via de regra, a avaliação de conformidade será feita com base no controlo interno, sendo realizada pelo fornecedor uma autoavaliação acerca do cumprimento de todos os requisitos obrigatórios e das normas técnicas aplicáveis, posteriormente consolidada em uma *documentação técnica* que ficará à disposição da autoridade de supervisão (os chamados organismos notificados) pelo prazo de 10 anos (art. 50.º). Nesta documentação técnica (art. 11.º e Anexo IV), deverão ser disponibilizadas aos organismos notificados todas as informações necessárias para aferir a conformidade do sistema de IA com esses requisitos, incluindo: uma descrição geral do sistema[12]; uma descrição pormenorizada dos elementos do sistema de IA e do respetivo processo de desenvolvimento[13]; informações pormenorizadas sobre

---

[12] Conforme descrito no Anexo IV, n. 1, neste ponto devem constar: a) A finalidade prevista, a(s) pessoa(s) responsáveis pelo seu desenvolvimento, a data e a versão do sistema; b) De que forma o sistema de IA interage ou pode ser utilizado para interagir com *hardware* ou *software* que não faça parte do próprio sistema de IA, se for caso disso; c) As versões do *software* ou *firmware* instalado e quaisquer requisitos relacionados com a atualização das versões; d) A descrição de todas as formas sob as quais o sistema de IA é colocado no mercado ou colocado em serviço; e) A descrição do *hardware* no qual se pretende executar o sistema de IA; f) Se o sistema de IA for um componente de produtos, fotografias ou ilustrações que revelem as características externas, a marcação e a disposição interna desses produtos; e g) Instruções de utilização para o utilizador e, se for caso disso, instruções de instalação.

[13] Segundo pontua o Anexo IV, n. 2, neste ponto devem constar: a) Os métodos utilizados e os passos dados com vista ao desenvolvimento do sistema de IA, incluindo, se for caso disso, o recurso a sistemas ou ferramentas previamente treinados

o acompanhamento, o funcionamento e o controlo do sistema de IA[14]; uma descrição pormenorizada do sistema de gestão de riscos; a descrição de todas as alterações introduzidas no sistema ao longo do seu ciclo

---

fornecidos por terceiros e de que forma estes foram utilizados, integrados ou modificados pelo fornecedor; b) As especificações de conceção do sistema, designadamente a lógica geral do sistema de IA e dos algoritmos; as principais opções de conceção, nomeadamente a lógica subjacente e os pressupostos utilizados, também no respeitante às pessoas ou grupos de pessoas em relação às quais se pretende que o sistema seja utilizado; as principais opções de classificação; o que se pretende otimizar com o sistema e a importância dos diferentes parâmetros; as decisões acerca de eventuais cedências em relação às soluções técnicas adotadas para cumprir os requisitos definidos no título III, capitulo 2; c) A descrição da arquitetura do sistema, explicando de que forma os componentes de *software* se incorporam ou enriquecem mutuamente e como se integram no processamento global; os recursos computacionais utilizados para desenvolver, treinar, testar e validar o sistema de IA; d) Se for caso disso, os requisitos de dados em termos de folhas de dados que descrevam as metodologias e técnicas de treino e os conjuntos de dados de treino utilizados, incluindo informações sobre a proveniência desses conjuntos de dados, o seu âmbito e as suas principais características; de que forma os dados foram obtidos e selecionados; procedimentos de rotulagem (por exemplo, para aprendizagem supervisionada), metodologias de limpeza de dados (por exemplo, deteção de valores atípicos); e) Análise das medidas de supervisão humana necessárias em conformidade com o artigo 14.o, incluindo uma análise das soluções técnicas necessárias para facilitar a interpretação dos resultados dos sistemas de IA pelos utilizadores, em conformidade com o artigo 13.o, n.o 3, alínea d); f) Se for caso disso, uma descrição pormenorizada das alterações predeterminadas do sistema de IA e do seu desempenho, juntamente com todas as informações pertinentes relacionadas com as soluções técnicas adotadas para assegurar a conformidade contínua do sistema de IA com os requisitos aplicáveis estabelecidos no título III, capítulo 2; g) Os procedimentos de validação e teste aplicados, incluindo informações sobre os dados de validação e teste utilizados e as principais características desses dados; as métricas utilizadas para aferir a exatidão, a solidez, a cibersegurança e a conformidade com outros requisitos aplicáveis estabelecidos no título III, capítulo 2, bem como potenciais impactos discriminatórios; registos dos testes e todos os relatórios de teste datados e assinados pelas pessoas responsáveis, incluindo no respeitante às alterações predeterminadas referidas na alínea f).

[14] Especialmente no que diz respeito às suas capacidades e limitações de desempenho, incluindo os níveis de exatidão no tocante a pessoas ou grupos de pessoas específicos em relação às quais se pretende que o sistema seja utilizado e o nível geral esperado de exatidão em relação à finalidade prevista; os resultados não pretendidos mas previsíveis e as fontes de riscos para a saúde e a segurança, os direitos fundamentais e a proteção contra a discriminação atendendo à finalidade prevista do sistema de IA; as medidas de supervisão humana necessárias em conformidade com o artigo 14.o, incluindo as soluções técnicas adotadas para facilitar a interpretação dos resultados dos sistemas de IA pelos utilizadores; especificações relativas aos dados de entrada, consoante apropriado (Anexo IV, n. 3).

de vida; uma lista das normas técnicas aplicadas total ou parcialmente; e uma descrição pormenorizada do sistema existente para avaliar o desempenho do algoritmo na fase de pós-comercialização.

De outro lado, porém, quando não for possível ao fornecedor aplicar total ou parcialmente normas harmonizadas desenvolvidas por organismos de normalização técnica ou quando estas normas não tiverem sido desenvolvidas, a avaliação de conformidade deverá ser baseada na análise do sistema de gestão de qualidade e da documentação técnica, caso em que, seguindo o procedimento detalhado no Anexo VII da *Proposta de Regulamento*, o processo de avaliação do cumprimento dos requisitos obrigatórios será realizado pelo organismo notificado que, ao final, vai emitir uma decisão fundamentada com as conclusões da sua avaliação acerca do sistema de inteligência artificial.

Importa mencionar que, para realizar esta avaliação, o organismo notificado deverá dispor de total acesso aos conjuntos de dados de treino e teste utilizados pelo fornecedor, incluindo através de interfaces de programação de aplicações ou outros meios e ferramentas adequadas que possibilitem o acesso remoto (Anexo VII, 4.3). Ao final da avaliação, o organismo notificado pode aprovar a documentação técnica emitindo um certificado UE de avaliação com validade máxima de 5 anos (prorrogável por iguais períodos mediante reavaliação, art. 44, n. 2), ou recusar a emissão do certificado UE informando o requerente do facto, fundamentando pormenorizadamente as razões da sua recusa. Em específico, caso a recusa se dê pelo não cumprimento dos requisitos relativos aos dados utilizados para treinar o sistema de IA, será necessário voltar a treinar o sistema de IA antes da apresentação do pedido de nova avaliação da conformidade (Anexo VII, n. 4.7).

Ainda como forma de aumentar a transparência e a supervisão públicas, de reforçar a supervisão pós-comercialização por parte das autoridades competentes e de garantir a confiança dos consumidores em produtos e serviços que utilizem algoritmos de inteligência artificial, após realizada a avaliação de conformidade necessária, os fornecedores deverão registar os sistemas de IA de risco elevado em uma base de dados unificada pública gerida pela Comissão Europeia na forma do artigo 60.º (Título VII). Finalmente, como forma de garantir a livre circulação dentro do mercado interno europeu, os sistemas de inteligência artificial de risco elevado devem apresentar, de modo visível, legível e indelével, a marcação CE para indicar o cumprimento de todos

os requisitos legais previstos na *Proposta de Regulamento* (Considerando 67 e artigo 49.º).

Tendo em consideração o caráter dinâmico e orgânico dos sistemas de inteligência artificial, que continuam a aprender depois de terem sido colocados no mercado ou em serviço, a tutela dos riscos inerentes aos sistemas de risco elevado não acaba em sua certificação técnica *ex ante*, motivo pelo qual, em seu Título VIII, a *Proposta de Regulamento* traz uma série de deveres de acompanhamento pós-comercialização na forma de obrigações de controlo e de comunicação aplicáveis aos fornecedores durante todo o ciclo de vida do algoritmo. Assim, em primeiro lugar, sob a perspetiva de tutelar eventuais riscos decorrentes de sistemas de IA de risco elevado, a *Proposta de Regulamento* obriga os fornecedores a informar as autoridades nacionais sobre incidentes graves ou anomalias que constituam violações do direito nacional e da União em matéria de direitos fundamentais assim que tomarem conhecimento das mesmas[15], bem como sobre eventuais recolhas ou retiradas de sistemas de IA do mercado (art. 62.º).

De outro lado, como forma de aproveitar a experiência adquirida na utilização de sistemas de IA de risco elevado, primeiro, para melhorar os algoritmos bem como os seus processos de conceção e desenvolvimento e, segundo, para assegurar uma resolução mais eficaz e atempada dos eventuais riscos observados após a colocação no mercado, a *Proposta de Regulamento* cria a obrigação para fornecedores de desenvolver e documentar um plano de acompanhamento pós-comercialização que seja proporcionado à natureza das tecnologias de IA e aos riscos do sistema (art. 61.º). Conforme detalham os artigos 3.º, n. 25 e 61.º do texto legislativo, o sistema de acompanhamento pós-comercialização deve recolher, documentar e analisar de forma ativa e sistemática os dados pertinentes fornecidos pelos utilizadores ou recolhidos por meio de outras fontes sobre o desempenho dos sistemas de IA de risco elevado ao longo da sua vida útil, bem como permitir ao fornecedor avaliar a contínua conformidade dos sistemas de IA com os requisitos estabelecidos no título III, capítulo 2.

---

[15] Conforme detalhado no artigo 62, n. 1, 2º parágrafo, "Essa notificação deve ser efetuada imediatamente após o fornecedor ter determinado uma relação causal entre o sistema de IA e o incidente ou anomalia ou a probabilidade razoável dessa relação e, em qualquer caso, o mais tardar 15 dias após o fornecedor ter conhecimento do incidente grave ou da anomalia".

7. Abordagem de caixa de areia e os ambientes de testagem da regulamentação

Neste contexto amplo da supervisão e certificação de sistemas de inteligência artificial de risco elevado, a Comissão Europeia destaca que, de um lado, as obrigações relativas à testagem *ex ante*, à gestão de riscos e à supervisão humana facilitarão o respeito de outros direitos fundamentais, graças à minimização do risco de decisões assistidas por IA erradas ou enviesadas em domínios críticos como a educação e a formação, o emprego, serviços essenciais, a manutenção da ordem pública e o sistema judicial. Complementarmente, os controlos e deveres de comportamento *ex post* inseridos no sistema de acompanhamento pós-comercialização funcionarão, para os consumidores europeus, como garantia de acesso a vias eficazes de recursos diante de eventuais violações a direitos fundamentais.

Ainda que assim seja, mesmo argumentando que "uma avaliação da conformidade *ex ante* abrangente por meio de controlos internos, aliada a uma forte execução *ex post*, poderá constituir uma solução eficaz e razoável para esses sistemas", a Comissão reconhece que o modelo de supervisão delineado na *Proposta de Regulamento* pode não ser o ideal para a tutela dos riscos conhecidos e desconhecidos oriundos da inteligência artificial, especialmente "dada a fase inicial da intervenção regulamentar e o facto de o setor da inteligência artificial ser bastante inovador e de só agora estarem a ser reunidos conhecimentos especializados para as auditorias".

Assim sendo, reconhecendo que a inteligência artificial é uma família de tecnologias em rápida evolução que exige novas formas de supervisão regulamentar e um espaço seguro para a experimentação, garantindo ao mesmo tempo uma inovação responsável e a integração de salvaguardas e medidas de atenuação dos riscos adequadas, a Comissão deixa uma margem de mudança e evolução às relações regulatórias através da adoção da chamada *abordagem de caixa de areia* (*sandbox approach*)[16] com a

---

[16] Esta abordagem de caixa de areia reflete o binómio excelência *vs* confiança estabelecido ainda no Livro Branco e pode ser encontrada também na revisão de 2021 do Plano Coordenado Para a Inteligência Artificial em que a Comissão esclarece que "Na sua essência, estes ambientes proporcionam uma instalação de experimentação para a regulamentação pública e permitem uma avaliação mais rápida do impacto da intervenção pública".

criação de *ambientes de testagem da regulamentação* que facilitem o desenvolvimento e o teste de sistemas de IA inovadores sob uma supervisão regulamentar rigorosa, antes que estes sistemas sejam colocados no mercado ou em serviço (Considerando 71).

Para isso, a Comissão incentiva os Estados-Membros a criar ambientes controlados para testar tecnologias inovadoras durante um período limitado com base num plano de testagem acordado com as autoridades competentes que reflita quatro objetivos principais, nomeadamente: (i) fomentar a inovação no domínio da IA, mediante a criação de um ambiente controlado de experimentação e teste na fase de desenvolvimento e pré-comercialização que assegure o cumprimento da legislação aplicável; (ii) reforçar a segurança jurídica para os inovadores; (iii) melhorar a supervisão e a compreensão, por parte das autoridades competentes, das oportunidades, dos riscos emergentes e dos impactos da utilização da inteligência artificial; e (iv) acelerar o acesso aos mercados, nomeadamente por via da eliminação dos entraves para as pequenas e médias empresas (PME) e as empresas em fase de arranque (Considerando 72).

8. Deveres de comportamento e elementos obrigatórios em sistemas de inteligência artificial de risco elevado

Como já se mencionou, partindo do entendimento de que os sistemas de inteligência artificial de risco elevado possuem um potencial de risco relevante para a saúde, a segurança e os direitos fundamentais dos consumidores europeus, a *Proposta de Regulamento* estabelece regras harmonizadas para o desenvolvimento, a colocação no mercado e a utilização de sistemas de IA na União Europeia, consolidando em seu Título III, capítulo 2, uma série de requisitos de segurança aplicáveis a este tipo de sistema de inteligência artificial. Em coerência com as recomendações e princípios internacionais[17] e como resultado de um trabalho preparatório de dois anos desenvolvido pelo Grupo de Peritos de Alto Nível sobre a Inteligência Artificial a *Proposta de Regulamento*

---

[17] Entre os documentos internacionais relevantes, cumpre citar os relatórios *The Age of Digital Interdependence*, publicado pela Organização das Nações Unidas, e *Artificial Intelligence in Society*, publicado pela Organização para a Cooperação e Desenvolvimento Económico (OCDE).

estabelece os requisitos legais aplicáveis aos sistemas de IA de risco elevado relativamente aos dados (isto é, à qualidade dos conjuntos de dados e à governação de dados), à documentação e à manutenção de registos, à transparência e à prestação de informações aos utilizadores, à supervisão humana, à solidez, à exatidão e à segurança.

Neste contexto, como forma de atenuar eficazmente os riscos e proteger os consumidores europeus, os fornecedores de sistemas de inteligência artificial de risco elevado devem criar e manter um sistema de gestão de qualidade[18] (art. 17.º) sólido e que documente, de forma sistemática e ordenada, políticas, procedimentos e instruções escritas para o cumprimento da legislação e dos requisitos técnicos aplicáveis aos algoritmos, em torno de seis elementos centrais, nomeadamente, procedimentos de gestão e governação de dados (art. 10), um sistema permanente de gestão de risco (art. 9), procedimentos de manutenção de registos (art. 12), garantia da transparência e da prestação de

---

[18] Conforme detalha o artigo 17.º da Proposta de Regulamento, o sistema de gestão da qualidade deve ser proporcionado à dimensão da organização do fornecedor e deve incluir, no mínimo, uma estratégia clara o cumprimento da regulamentação, incluindo procedimentos de avaliação da conformidade e de gestão de modificações do sistema de IA de risco elevado; descrição das técnicas, procedimentos e ações sistemáticas que serão utilizadas para o desenvolvimento, o controlo da qualidade e a garantia da qualidade sistema e para a conceção, o controlo da conceção e a verificação da conceção do sistema a terceiros; procedimentos de exame, teste e validação a realizar antes, durante e após o desenvolvimento sistema e a frequência com a qual serão realizados; a descrição dos testes e procedimentos que devem ser feitos em ambientes controlados que facilitem o desenvolvimento, a testagem e a validação dos sistemas, sob supervisão e orientação das autoridades competentes (art. 53); as especificações técnicas do sistema, incluindo normas técnicas a aplicar, meios a usar para assegurar que o sistema cumpre os requisitos, se as normas harmonizadas em causa não forem aplicadas na íntegra; sistemas e procedimentos de gestão de dados, incluindo sua recolha, análise, rotulagem, armazenamento, filtragem, prospecção, agregação, conservação e quaisquer outras operações relativa aos dados que realizadas antes e para efeitos da colocação no mercado ou colocação em serviço; o sistema de gestão de riscos (artigo 9.º); estabelecimento, aplicação e manutenção de um sistema de acompanhamento pós-comercialização (art. 61.º); procedimentos de comunicação de incidentes graves e de anomalias (art. 62.º); a gestão da comunicação com autoridades nacionais competentes, incluindo as autoridades setoriais, organismos notificados, outros operadores, clientes ou outras partes interessadas; sistemas e procedimentos de manutenção de registos de toda a documentação e informação importante; gestão de recursos, incluindo medidas relacionadas com a segurança do aprovisionamento; um quadro que defina as responsabilidades do pessoal com funções de gestão e do restante pessoal no atinente a todos os aspetos elencados no presente número.

informações (art. 13), mecanismos de supervisão humana (art. 14) e garantia de exatidão, solidez e cibersegurança (art. 15).

Diante da importância da disponibilidade de dados de elevada qualidade para o desempenho de vários sistemas de IA com vista a assegurar que o sistema de IA de risco elevado funcione como pretendido e de modo seguro e não se torne a fonte de uma discriminação proibida pelo direito da União, sobretudo quando são utilizadas técnicas que envolvem o treino de modelos, o primeiro elemento relevante para o sistema de gestão de qualidade pode ser identificado nos procedimentos de governação de dados (Considerando 44 e art. 10). Em casos de modelos de inteligência artificial treinados com dados, os fornecedores devem garantir não apenas a fiabilidade e a solidez dos procedimentos de recolha, pré-processamento e processamento dos dados[19], mas também devem garantir a qualidade dos conjuntos de dados de treino, validação e teste, assegurando que são pertinentes, representativos, isentos de erros e completos; que têm as propriedades estatísticas adequadas no tocante às pessoas ou grupos de pessoas em que o sistema será utilizado; e que têm em conta as características, os elementos idiossincráticos do enquadramento geográfico, comportamental e funcional em que o algoritmo deve ser utilizado.

Um segundo elemento essencial aos sistemas de gestão de qualidade de sistemas de inteligência artificial de risco elevado são os procedimentos de manutenção de registos, de que trata o artigo 12.º da *Proposta de Regulamento*. Considerada essencial para verificar o cumprimento dos requisitos estabelecidos na *Proposta de Regulamento* e para a identificação e resolução e atempada de eventuais riscos identificados após a colocação no mercado, a manutenção de registos deve incluir informações relevantes como as características gerais, as capacidades e as limitações do sistema, os algoritmos, os dados e os processos de

---

[19] Conforme descreve o artigo 10.º, os fornecedores devem garantir que os conjuntos de dados de treino, validação e teste estão sujeitos a práticas adequadas de governação e gestão de dados relacionados às escolhas de conceção tomadas; aos procedimentos de recolha de dados; ao pré-processamento dos dados, isto é, à preparação e o tratamento de dados, tais como anotação, rotulagem, limpeza, enriquecimento e agregação dos dados; à formulação dos pressupostos aplicáveis, nomeadamente informações que os dados devem medir e representar; à avaliação prévia da disponibilidade, quantidade e adequação dos conjuntos de dados que são necessários; aos exames para detetar eventuais enviesamentos; à identificação de eventuais lacunas ou deficiências de dados e de possíveis soluções para as mesmas.

treino, teste e validação utilizados, bem como documentação sobre o sistema de gestão de riscos aplicado (Considerando 46).

Neste contexto, será obrigatório o desenvolvimento de um sistema de registo automático de eventos enquanto o algoritmo estiver em funcionamento, registos esses que deverão ser mantidos por um período adequado em função da finalidade prevista do sistema e das obrigações legais aplicáveis. Ademais, deverá ser assegurado um nível de rastreabilidade do funcionamento adequado à finalidade prevista do sistema, além de ser assegurado que as capacidades de registo permitirão o controlo do funcionamento do sistema contra riscos para a saúde e segurança ou para direitos fundamentais, contra modificações substanciais e para facilitar o acompanhamento pós-comercialização do algoritmo.

Outro elemento essencial para a gestão dos sistemas de inteligência artificial de risco elevado é o sistema permanente de gestão de riscos (art. 9.º), que tem como objetivo fazer com que eventuais riscos residuais associados a cada perigo e eventuais riscos globais associados ao algoritmo possam ser considerados aceitáveis, quando do uso normal do sistema ou quando houver um uso indevido razoavelmente previsível. Tendo este objetivo em mente, o sistema permanente de gestão de riscos deverá consistir num processo iterativo contínuo, executado ao longo de todo o ciclo de vida de um sistema de IA de risco elevado e regularmente submetido a atualizações sistemáticas, dividido em quatro etapas: (i) identificação e análise dos riscos conhecidos e previsíveis associados a cada sistema de IA de risco elevado; (ii) estimativa e avaliação de riscos que podem surgir quando o sistema de IA de risco elevado é usado em conformidade com a sua finalidade prevista e em condições de utilização indevida razoavelmente previsíveis; (iii) avaliação de outros riscos que possam surgir, baseada na análise dos dados recolhidos a partir do sistema de acompanhamento pós-comercialização; e (iv) adoção de medidas de gestão de riscos adequadas levando consideração, primeiro, os efeitos e eventuais interações resultantes da aplicação combinada dos requisitos obrigatórios e, segundo, o estado da técnica geralmente reconhecido em normas harmonizadas ou especificações comuns pertinentes.

Reconhecendo que os sistemas de inteligência artificial podem ser caracterizados por uma opacidade que os pode tornar incompreensíveis ou demasiado complexos para as pessoas singulares, a *Proposta de Regulamento* torna também obrigatória a garantia da transparência e a

prestação de informações para os utilizadores e consumidores europeus de forma a garantir que sejam capazes de interpretar o resultado do sistema e utilizá-lo de forma adequada (Considerando 47). Com isso, por força do artigo 13 do texto legislativo, os sistemas de IA de risco elevado devem ser acompanhados de documentação pertinente e instruções de utilização, em formato digital ou outro adequado, além de incluir informações concisas, completas, claras e compreensíveis relativas a possíveis riscos para os direitos fundamentais e de discriminação, se for caso disso[20].

Outra forma importante de tutela do risco inerente aos sistemas de IA previsto exigido pela Comissão Europeia são os mecanismos de supervisão humana (art. 14.º), em decorrência dos quais os sistemas de IA de risco elevado devem ser concebidos e desenvolvidos de maneira que permita a sua supervisão por pessoas singulares de forma a prevenir ou minimizar os riscos para a saúde, a segurança ou os direitos fundamentais que possam surgir quando um sistema de IA de risco elevado é usado em conformidade com a sua finalidade prevista ou em condições de utilização indevida razoavelmente previsí-

---

[20] Conforme detalha o n. 3 do artigo 13.º, as informações obrigatórias devem especificar:

a) A identidade e os dados de contacto do fornecedor e, se for caso disso, do seu mandatário; b) As características, capacidades e limitações de desempenho do sistema de IA de risco elevado, incluindo: i) a finalidade prevista do sistema, ii) o nível de exatidão, solidez e cibersegurança a que se refere o artigo 15.o relativamente ao qual o sistema de IA de risco elevado foi testado e validado e que pode ser esperado, bem como quaisquer circunstâncias conhecidas e previsíveis que possam ter um impacto nesse nível esperado de exatidão, solidez e cibersegurança, iii) qualquer circunstância conhecida ou previsível, relacionada com a utilização do sistema de IA de risco elevado de acordo com a sua finalidade prevista ou em condições de utilização indevida razoavelmente previsíveis, que possa causar riscos para a saúde e a segurança ou os direitos fundamentais, iv) o desempenho do sistema no tocante às pessoas ou grupos de pessoas em que o sistema se destina a ser utilizado, v) quando oportuno, especificações para os dados de entrada, ou quaisquer outras informações importantes em termos dos conjuntos de dados de treino, validação e teste usados, tendo em conta a finalidade prevista do sistema de IA; c) As alterações do sistema de IA de risco elevado e do seu desempenho que foram predeterminadas pelo fornecedor aquando da avaliação da conformidade inicial, se for caso disso; d) As medidas de supervisão humana a que se refere o artigo 14.o, incluindo as soluções técnicas adotadas para facilitar a interpretação dos resultados dos sistemas de IA pelos utilizadores; e) A vida útil esperada do sistema de IA de risco elevado e quaisquer medidas de manutenção e assistência necessárias para assegurar o correto funcionamento desse sistema de IA, incluindo no tocante a atualizações do software.

veis (Considerando 48). Neste ponto, a *Proposta de Regulamento* torna obrigatória a integração aos sistemas de IA de restrições operacionais que não possam ser neutralizadas pelo próprio sistema, compelindo os sistemas a responder ao operador humano, além de tornar obrigatório que as pessoas singulares a quem foi atribuída a supervisão humana tenham as competências, a formação e a autoridade necessárias para desempenhar essa função[21].

Por fim, o último elemento essencial previsto pela *Proposta de Regulamento* para a tutela dos riscos inerentes aos sistemas de inteligência artificial de risco elevado é a garantia da exatidão, solidez e cibersegurança do algoritmo. Conforme explica a Comissão Europeia nos Considerandos 49 a 50, é importante garantir não apenas que os sistemas de IA tenham um desempenho coerente ao longo de todo o seu ciclo de vida mas, igualmente, que estes sistemas sejam resistentes aos riscos associados às suas limitações inerentes – como erros, falhas, incoerências ou situações inesperadas – e às ações maliciosas suscetíveis de pôr em causa a segurança dos sistema de IA ou mesmo dar origem a comportamentos prejudiciais indesejáveis. Da mesma forma, é essencial garantir que os sistemas de IA sejam resistentes às ações de terceiros mal-intencionados que venham a tentar explorar as vulnerabilidades dos sistemas com o objetivo de lhes alterar a utilização, o comportamento e o desempenho ou por em causa as propriedades de segurança, explorando vulnerabilidades dos ativos digitais do sistema de IA ou

---

[21] Em específico, a *Proposta de Regulamento* torna obrigatória a adoção de medidas, pelo fornecedor, para garantir que as pessoas responsáveis pela supervisão humana dos sistemas de IA de risco elevado (i) Compreendam completamente as capacidades e limitações do sistema de IA de risco elevado e sejam capazes de controlar devidamente o seu funcionamento, de modo que os sinais de anomalias, disfuncionalidades e desempenho inesperado possam ser detetados e resolvidos o mais rapidamente possível; (ii) estejam conscientes da possível tendência para confiar automaticamente ou confiar excessivamente no resultado produzido pelo sistema de IA de risco elevado («enviesamento da automatização»), em especial relativamente aos sistemas de IA de risco elevado usados para fornecer informações ou recomendações com vista à tomada de decisões por pessoas singulares; (iii) sejam capazes de interpretar corretamente o resultado do sistema de IA de risco elevado, tendo em conta, nomeadamente, as características do sistema e as ferramentas e os métodos de interpretação disponíveis; (iv) em qualquer situação, sejam capazes de decidir não usar o sistema de IA de risco elevado ou ignorar, anular ou reverter o resultado do sistema de IA de risco elevado; (v) sejam capazes de intervir no funcionamento do sistema ou interromper o sistema por meio de um botão de «paragem» ou procedimento similar.

da infraestrutura de tecnologias da informação e comunicação (TIC) subjacente.

Por este motivo, os fornecedores devem garantir um nível apropriado de exatidão dos sistemas, informando aos utilizadores as métricas utilizadas para esta avaliação, além de garantir a solidez técnica e a cibersegurança dos algoritmos de acordo com o estado da técnica geralmente reconhecido, evitando a ocorrência de problemas de segurança possam afetar negativamente os direitos fundamentais, por exemplo, devido a decisões erradas ou a resultados errados ou enviesados gerados pelo sistema de IA. Para alcançar estes objetivos, o artigo 15.º da *Proposta de Regulamento* prevê que, tendo em conta a sua finalidade prevista, os sistemas de IA de risco elevado devem ser concebidos com um nível adequado de exatidão, solidez e cibersegurança, sendo resistentes a erros, falhas ou incoerências que possam ocorrer no sistema ou no ambiente em que aquele opera, em especial devido à interação com pessoas singulares ou outros sistemas; e sendo também resistentes a tentativas de terceiros não autorizados de alterar a sua utilização ou desempenho explorando as vulnerabilidades do sistema.

9. Conclusão

A *Proposta de Regulamento* sobre inteligência artificial vem consolidar os princípios e objetivos delineados em documentos anteriores sobre IA e, neste contexto, cria um quadro normativo básico relativo à governação, à supervisão e à responsabilidade em sistemas de inteligência artificial, estabelecendo regras harmonizadas para o desenvolvimento, a colocação no mercado e a utilização de sistemas de IA na União Europeia e criando requisitos essenciais e obrigatórios para determinados tipos de sistemas de IA e de deveres de comportamento relacionados à transparência e à prestação de informações.

Como reflexo da busca pelo equilíbrio regulatório e em consolidação da abordagem europeia centrada no binómio excelência e confiança descrita no *Livro Branco*, a *Proposta de Regulamento* busca criar uma intervenção jurídica equilibrada e proporcional. Para isso, propõe um quadro jurídico sólido centrado em uma *abordagem baseada no risco* que classifica os sistemas de inteligência artificial a partir dos níveis de risco criados pelos sistemas e correlaciona deveres de comportamento específicos e proporcionais a cada tipo, especificamente: *riscos*

inaceitáveis (Título II), que são práticas proibidas em território europeu; *riscos limitados* (Título IV), para os quais existem deveres de informação e transparência para com consumidores; *riscos mínimos* (Título IX), para os quais não existem deveres e obrigações específicos; e, *riscos elevados* (Título III), para os quais são previstas uma série de requisitos relacionados à qualidade dos dados, à documentação e à rastreabilidade, à transparência, à supervisão humana, à exatidão e à solidez.

Para avaliar o cumprimento dos requisitos obrigatórios e certificar os sistemas de inteligência artificial de risco elevado, a *Proposta de Regulamento* segue o modelo já utilizado no âmbito do *novo quadro legislativo*, criando relações regulatórias *ex ante*, relacionados à testagem, prestação de informações e documentação *antes* da colocação de um sistema de IA no mercado; e *ex post,* de controlo, manutenção de registos e prestação de informações sobre incidentes graves ou anomalias no *pós-comercialização*, durante todo o ciclo de vida do algoritmo.

Referências bibliográficas

Ebers, Martin, «Standardizing AI – The Case of the European Commission's Proposal for an Artificial Intelligence Act», in Larry A. DiMatteo / Michel Cannarsa / Cristina Poncibò, ed., *The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics*, Cambridge University Press 2022.

Godinho, Inês Fernandes / Flores, Cláudio R. / Marques, Nuno Castro, «Consultation on The White Paper on Artificial Intelligence – A European Approach», *ILP Law Review / Revista de Direito da ULP* 14/1 (2021) 157-167, doi:10.46294.

Lilkov, Dimitar, «Regulating artificial intelligence in the EU: A risky game», *European View* 20/2 (2021) 166–174.

Ramos, José Ricardo Marcondes, «Relatório sobre a atual regulação normativa europeia e portuguesa em matéria de Inteligência Artificial», *Revista Portuguesa de Ciência Criminal* 31 (2022) 633-646.

Veale, Michael / Borgesius, Frederik Zuiderveen, «Demystifying the Draft EU Artificial Intelligence Act», *Computer Law Review International* 22/4 (2021) 97–112.

4

Sistemas de Armas Autónomas e Respectiva Regulamentação

(https://doi.org/10.47907/DireitoemMudanca/2023/4)

*Miguel João Costa**

Resumo: No presente texto faz-se uma breve análise da Proposta de Regulamento da União Europeia (UE) sobre Inteligência Artificial (IA) na parte respeitante à matéria militar e de defesa e à matéria da segurança nacional, que se encontram expressamente excluídas do seu âmbito de aplicação. Procura-se compreender as razões dessa exclusão, reflectir sobre se ela é de facto total e apurar a que regras ficam então sujeitas aquelas matérias. Conclui-se que a exclusão – que é menos censurável quanto à matéria militar e de defesa do que quanto à de segurança nacional, embora seja em ambos os casos inquietante – se deve essencialmente a razões de soberania e estratégia; que o Regulamento poderá, ainda assim, produzir benefícios colaterais quanto às matérias excluídas, embora seja optimista supô-los significativos; e que a regulação do recurso à IA no domínio militar permanece confiada essencialmente ao Direito da Guerra, que não tem vocação para lhe impor limites consideráveis.

Palavras-chave: Armas Autónomas Mortíferas; Direito da Guerra; Direito Penal Internacional; Inteligência Artificial; Regulamento da União Europeia

* Professor Auxiliar, Univ Coimbra, IJ, FDUC.

Por opção do autor, o presente texto é escrito segundo o antigo Acordo Ortográfico da Língua Portuguesa.

Introdução

Gostaria de começar por agradecer ao Instituto Jurídico e à Organização deste evento – à Senhora Doutora Susana Aires de Sousa e à Senhora Doutora Maria João Antunes, coordenadora da linha de investigação em que ele se integra, a ambas com grande amizade e admiração académica – o convite para nele participar, que me permite dar expressão ao trabalho que tenho vindo a fazer neste âmbito e expandi-lo um pouco em termos temáticos.

A minha área de trabalho é o Direito Penal e a investigação (até ver apenas exploratória) que tenho feito a propósito da IA tem incidido sobre questões transnacionais e internacionais, geralmente em co-autoria com outros investigadores – como é o caso, quanto à temática militar, de Miguel Lemos[1]. Apresento aqui algumas reflexões sobre esta temática no plano do Direito da UE, que ainda não tinha analisado.

1. A Proposta de Regulamento da UE

A Proposta de Regulamento, desde a sua versão original, exclui do seu âmbito de aplicação os "sistemas de IA desenvolvidos ou utilizados exclusivamente para efeitos militares" – considerando 12[2], confirmado no artigo 2.º, n.º 3, nestes termos: "O presente regulamento não se aplica aos sistemas de IA desenvolvidos ou usados exclusivamente para fins militares."[3]

Sendo o propósito deste evento discutir a Proposta de Regulamento e estando a matéria militar (e, agora, também outras matérias adjacentes, como se verá) excluída dessa Proposta, a primeira e muito

---

[1] Cf. esp. Miguel Lemos / Miguel João Costa, «Inteligência Artificial e Direito da Guerra: Reflexões sobre as Armas Autónomas Mortíferas», in Anabela Miranda Rodrigues, coord., *A Inteligência Artificial no Direito Penal: Volume II*, Coimbra: Almedina, 2022, 91-123. Sobre outras questões, mas ainda nos planos internacional e transnacional, cf. também Miguel João Costa / António Manuel Abrantes, «Os Desafios da Inteligência Artificial da Perspectiva Transnacional: A Jurisdição e a Cooperação Judiciária», in Anabela Miranda Rodrigues, coord., *A Inteligência Artificial no Direito Penal*, Coimbra: Almedina, 2020, 163-217.

[2] Bruxelas, 21.4.2021 COM(2021) 206 final 2021/0106 (COD). Cf., na última versão, o considerando 12-A: Bruxelas, 6.12.2022 (OR. en) 15698/22, 2021/0106(COD).

[3] Na versão inicial – depois alterada também neste ponto, como se verá.

legítima questão que importa enfrentar é a de saber em que medida se trata aqui realmente de um tema. Creio, porém, que de um tema se trata sim. A própria circunstância de a Proposta excluir expressamente aquela matéria do seu âmbito de aplicação já denota que lhe atribui uma importância particular; que essa matéria se considerava uma candidata natural à regulação e a que sua exclusão, por isso, carecia de ser explicitada. O que suscita, pelos menos, as três seguintes questões:

1.1 Porque é que essa matéria foi excluída?

1.2 Essa exclusão é mesmo total?

1.3 Onde se regula então aquela matéria?

1.1 Na versão original da Proposta excluíam-se apenas os *fins militares* e não se fazia qualquer especificação quanto à natureza das entidades em causa. Por outro lado, no considerando 12 fazia-se referência ao Título V do Tratado da União Europeia (TUE), relativo à Acção Externa da União. Esse Título, no entanto, não especifica o que deva entender-se por "fins militares". Refere "meios militares" e outras expressões próximas[4], mas não define "fins militares" e, em vários preceitos, refere-se ao "domínio militar *ou* [de modo alternativo, portanto] da defesa"[5]. Isso sugere que o "domínio militar" e o "domínio da defesa" terão alguma autonomia, por um lado; mas, do mesmo passo – e uma vez que surgem sempre acoplados –, que se entende deverem receber enquadramento jurídico idêntico[6].

A Presidência Eslovena do Conselho da UE, em Novembro de 2021, propôs – entre outras coisas – que o Regulamento: (i) Justificasse a exclusão da matéria militar; (ii) Esclarecesse no considerando 12 que, se o sistema de IA, embora desenvolvido para fins exclusivamente militares, fosse utilizado para outros fins, o Regulamento se lhe deveria aplicar; e (iii) Excluísse do seu âmbito de aplicação, além da matéria militar, também a da segurança nacional[7].

Em aberto permaneciam ainda algumas questões, designadamente a de saber se o Regulamento abrangeria sistemas de IA desenvolvidos por entidades privadas por sua própria iniciativa, mas com vista a

---

[4] *V.g.*, "capacidades militares": cf. os artigos 41.º, n.º 3, § 2, 42.º, n.º 6, 45.º, n.º 1, alíneas a) e c), e 46.º, n.º 1.

[5] Cf. os artigos 31.º, n.º 4, 41.º, n.º 2 e 48.º, n.º 7.

[6] Sobre isto, cf. brevemente *infra*, em texto.

[7] Bruxelas, 29.11.2021 (OR. en) 14278/21, 2021/0106(COD), 3, 11 e 32.

serem disponibilizados aos Estados-Membros para utilização com fins exclusivamente militares. Também a de saber se certas actividades militares e de defesa, para efeitos do Regulamento, incluiriam acções que não envolvem tipicamente a utilização de força letal, designadamente "de desarmamento, as missões humanitárias e de evacuação (...), de prevenção de conflitos e de manutenção de paz"[8]. Por exemplo, a utilização, numa acção de manutenção de paz ou numa missão humanitária, de sistemas de identificação biométrica à distância com funcionalidade de reconhecimento do estado emocional[9].

Percorridas outras etapas do processo legislativo, a versão mais recente da Proposta[10]: clarifica que a exclusão vale para entidades tanto públicas como privadas; alarga a exclusão (se é que não se pretendia que isso decorresse já da versão inicial da Proposta, lida em conjugação com o TUE) à defesa e à segurança nacional; e fundamenta a exclusão de modo mais detalhado, o que faz, repare-se, agrupando os fins *militares* e de *defesa* (em consonância com o que acontece no âmbito do TUE, nos termos acima indicados) e autonomizando-os dos fins de *segurança nacional* (embora a exclusão de uns e de outros partilhe alguns fundamentos, na medida em que ambos se reconduzem a funções essenciais do Estado que a União está vinculada a respeitar):

- Quanto aos *fins militares e de defesa*, a exclusão é justificada: por um lado, pelo disposto no artigo 4.º, n.º 2, do TUE (segundo o qual, para o que aqui mais releva, a UE "respeita as funções essenciais do Estado, nomeadamente as que se destinam a *garantir a integridade territorial*"); por outro, pelas "especificidades da política de defesa dos Estados-Membros e da União" (abrangidas pelo referido Título V, capítulo 2, do TUE)[11], as quais "estão sujeitas ao direito internacional público, que é, por conseguinte, o quadro jurídico mais adequado para a regulamentação dos sistemas de IA no contexto da utilização da força letal e de outros sistemas de IA no contexto de atividades militares e de defesa"[12].

---

[8] Previstas nos artigos 42.º, n.º 1, e 43.º, n.º 1, do TUE, como acções em que a UE pode usar meios militares.

[9] Cf. *v.g.* os pontos 6, alínea b), e 7, alínea a), do Anexo III da versão final da Proposta de Regulamento, anexo onde se especifica os sistemas de IA "de risco elevado" para os efeitos do respectivo artigo 6.º, n.º 3.

[10] A já referida versão de 6.12.2022 (OR. en) 15698/22, 2021/0106(COD).

[11] *Id.*, Considerando 12-A.

[12] *Ibid.*

- Quanto aos *fins de segurança nacional*, a exclusão é justificada: por um lado, novamente pelo artigo 4.º, n.º 2, do TUE (agora, se bem vejo, na dimensão onde se prevê que a União "respeita as funções essenciais do Estado, nomeadamente as que se destinam (...) a manter a *ordem pública* e a salvaguardar a *segurança nacional*", dimensão relativamente à qual aquele preceito acrescenta a asserção, que não tem paralelo quanto à matéria militar e de defesa, de que "a segurança nacional continua a ser da exclusiva responsabilidade de cada Estado-Membro"); por outro, pela "natureza específica e pelas necessidades operacionais específicas das atividades de segurança nacional e pelas regras nacionais específicas aplicáveis a essas atividades".

Isto ajuda a responder à questão de saber porque é que estas matérias foram excluídas do âmbito de aplicação da Proposta de Regulamento.

1.2 A última versão da Proposta esclarece também que um sistema de IA, tendo embora sido desenvolvido, colocado no mercado, colocado em serviço ou utilizado para fins militares, de defesa ou de segurança nacional, for temporária ou permanentemente utilizado para fins não excluídos (*v.g.*, civis ou humanitários, de manutenção da ordem pública ou de segurança pública), então será abrangido pelo âmbito de aplicação do Regulamento, devendo as entidades que o utilizarem para estes fins assegurar a sua conformidade com o Regulamento, a menos que já esteja em conformidade com ele[13]. Esclarece ainda, em termos idênticos, que um sistema de IA colocado no mercado ou em serviço para um fim excluído (*sc.*, militar, de defesa ou de segurança nacional) e, simultaneamente, para um ou mais fins não excluídos (*v.g.*, civis), então integrará aquele âmbito de aplicação, devendo a sua conformidade com o Regulamento ser assegurada pelos respectivos fornecedores.

Isto esclarece algumas das dúvidas colocadas anteriormente e ajuda a responder à questão de saber se a exclusão das matérias militares, de defesa e de segurança nacional é mesmo completa: talvez possa dizer-se que não o é, embora num sentido muito específico e muito limitado[14], uma vez que, mesmo naqueles casos, a *utilização* do sistema de IA pelas entidades que realizam atividades de segurança nacional, de defesa e

---

[13] *Ibid.*, também para o que segue.

[14] Sobre isto, cf. ainda *infra*, neste mesmo ponto.

militares, independentemente do tipo de entidade em causa, não fica sujeita, quando destinada a esses fins, à aplicação do Regulamento. O que vale, de igual modo, para sistemas colocados no mercado para fins civis ou de manutenção da ordem pública, quando sejam utilizados – com ou sem modificação – para fins militares, de defesa ou de segurança nacional.

Antes de regressar a esse ponto relativo à amplitude da exclusão, justifica-se uma breve reflexão crítica sobre a *exclusão da matéria da segurança nacional*. A matéria militar e a matéria da defesa, independentemente das diferenças que haja entre elas, são, enquanto conjunto (e sobretudo se enquanto tal forem juridicamente enquadradas[15]), relativamente fáceis de autonomizar de outras matérias que de algum modo se aproximam delas, como a da investigação e repressão criminais. Já a fronteira entre a segurança nacional e estas outras matérias é muito menos nítida – e é-o cada vez menos. Bastará pensar na evolução da política criminal e da dogmática penal no domínio do terrorismo e em outros domínios atraídos ao conceito do "direito penal do inimigo" – que, de resto, tende a contaminar mesmo domínios para que não foi directamente pensado e que, apesar do amplo repúdio académico, continua a florescer na hipersensibilidade das sociedades contemporâneas ao risco e à insegurança[16]. Isso pode levantar problemas quanto a determinar se certas utilizações do sistema de IA se reconduzem a um âmbito ou a outro. Além disso, embora o Tribunal de Justiça da União Europeia já tenha sustentado que a responsabilidade exclusiva para a segurança nacional referida no n.º 2 do artigo 4.º do TUE tem os seus limites[17], a autonomia de que os Estados-Membros aí dispõem é

---

[15] Como no TUE: cf. *supra*.

[16] Cf. Mariona Llobet, "Enemy Criminal Law (*Feindstrafrecht*)", in Pedro Caeiro *et al.*, ed., *Elgar Encyclopaedia of Crime & Criminal Justice*, Cheltenham: Edward Elgar Publishing, 2023 (https://www.elgaronline.com/edcollchap/book/9781789902990/b-9781789902990.enemy.criminal.law.xml).

[17] *V.g.*, em *La Quadrature du Net* (C-511/18 *et al.*), de 6 de Outubro de 2020, § 99: "(...) Com efeito, em conformidade com jurisprudência constante do Tribunal de Justiça, embora incumba aos Estados-Membros definir os seus interesses essenciais de segurança e adotar as medidas adequadas para garantir a sua segurança interna e externa, o simples facto de uma medida nacional ter sido adotada para efeitos da proteção da segurança nacional não pode implicar a inaplicabilidade do direito da União e dispensar os Estados-Membros do respeito necessário desse direito" (com ulteriores referências jurisprudenciais).

muito considerável, o que pode provocar um efeito de transbordamento ('*spill-over*') negativo, no sentido de um sistema de IA desenvolvido para fins de segurança nacional – isento, portanto, das exigências do Regulamento – acabar por migrar, com pouco ou nenhum controlo, para áreas a que ele deveria aplicar-se, como a da investigação e repressão criminais 'normais'[18].

Quanto às matérias militares e de defesa propriamente ditas, a sua exclusão apresenta-se como qualquer coisa de inexorável. Não é plausível que se desenvolva no Direito da UE um corpo significativo de normas paralelo ao do direito internacional público a que pudesse chamar-se de um 'Direito da Guerra da UE'[19]. A regulação destas matérias tem uma vocação internacional, uma ambição de universalidade que permita constranger não só a conduta própria em matéria de guerra, mas também a das forças inimigas. Há aqui a preocupação de assegurar uma simetria de forças em relação a outros blocos de Estados, preocupação que não convida a que determinado Estado ou conjunto de Estados se auto-imponha parâmetros de actuação ético-jurídicos muito mais elevados do que aqueles que vinculam os demais.

De todo o modo, há quem consiga ver aqui um copo meio cheio, no sentido de destacar que as matérias excluídas do âmbito de aplicação do Regulamento, tanto a militar e de defesa como a de segurança nacional, poderão beneficiar de um efeito de transbordamento agora *positivo*: em virtude da aplicação do Regulamento às outras matérias e aos sistemas de IA de utilização dupla[20], os sistemas de IA

---

[18] Cf. European Center for Not-for-Profit Law, "Artificial Intelligence Act Amendments: Scope of the EU Artificial Intelligence Act (AIA): Military Purposes and National Security", 2022 (aqui), ilustrando este risco com o caso da tecnologia de espionagem *Pegasus* (embora não creia que este *spyware*, em si mesmo, envolva a utilização de um sistema de IA: cf. mais aqui), desenvolvida por uma empresa israelita (e já utilizada também em Estados-Membros da UE: cf. também aqui, 124 ss.); esta tecnologia permite a infiltração em telemóveis e outros dispositivos de modo remoto e sem qualquer interacção por parte do utilizador ('*zero-click attacks*') e, embora desenvolvida com propósitos de segurança nacional, terá rapidamente passado a ser utilizada para outros fins.

[19] Cf. ainda *infra*, II.

[20] Com relevo, cf. ainda o Regulamento (UE) 2021/821 do Parlamento Europeu e do Conselho, de 20 de Maio de 2021, que cria um regime de controlo das exportações, corretagem, assistência técnica, trânsito e transferências de produtos de dupla utilização (aqui), cujo artigo 2.º, n.º 1, define estes produtos como aqueles "que possam ser utilizados tanto para fins civis como para fins militares, incluindo

desenvolvidos sem exclusivos propósitos militares e de defesa ou de segurança nacional, ainda que venham depois a ser utilizados para esses propósitos, pelo menos terão em certo momento sido submetidos ao crivo do Regulamento[21]. É neste sentido específico – muito específico – que talvez possa dizer-se que a matéria militar e de defesa e a matéria da segurança nacional não ficam completamente intocadas por este empreendimento regulatório.

1.3 A sua regulação propriamente dita, contudo, passará por outros planos normativos. No caso da matéria militar e de defesa, fundamentalmente, pelo Direito Internacional Público, mais especificamente o Direito Internacional Humanitário, Direito da Guerra ou *Jus in Bello*.

O Parlamento Europeu, numa Resolução de 20 de Janeiro de 2021 sobre "inteligência artificial: questões de interpretação e de aplicação do direito internacional na medida em que a UE é afetada nos domínios da utilização civil e militar e da autoridade do Estado fora do âmbito da justiça penal"[22], manifestou forte preocupação com o desenvolvimento da IA no domínio militar: sublinhou que a utilização militar (também a civil) da IA deve "imperativamente" manter-se sob "controlo humano significativo"[23]; reforçou a ideia de que na conceção, desenvolvimento, ensaio, implantação e utilização de sistemas de IA para estes fins "é imperativo ter em devida conta os riscos potenciais em qualquer altura"[24]; insistiu na necessidade de uma estratégia da União "contra" as armas autónomas letais e na "proibição" de robôs assassinos.[25] Em última instância, porém, não excluiu totalmente que, embora como "*último recurso*"[26] e dentro de condições exigentes[27], aquelas armas sejam utilizadas.

---

produtos que possam ser utilizados na conceção, desenvolvimento, produção ou utilização de armas nucleares, químicas ou biológicas e dos seus meios de lançamento, incluindo todos os produtos que possam ser utilizados tanto para fins não explosivos como para de qualquer modo auxiliar no fabrico de armas nucleares ou outros engenhos explosivos nucleares".

[21] Cf. Ronja Rönnback, "Challenges of Governing AI for Military Purposes and Spill-Over Effects of the AI Act", *The AI Alliance Blog*, 27 February 2023 (aqui).

[22] P9_TA(2021)0009 (aqui).

[23] *Id.*, §§ 3, 49.

[24] *Id.*, § 11.

[25] *Id.*, § 29; cf. todos os §§ 27-50.

[26] *Id.*, § 34.

[27] Cf. *v.g.* o § 37.

Lamentou a inércia no desenvolvimento de instrumentos jurídicos internacionais sobre estas armas e defendeu que a União deve assumir um "papel activo" e de "liderança" nesse processo[28]. Em última instância, porém, parece conceder que a regulação destas matérias está deixada ao Direito da Guerra – e que, portanto, pelo menos para já, apenas poderá encontrar limites nos parâmetros dele decorrentes, designadamente das Convenções de Genebra, dos ditames de humanidade e consciência pública da Cláusula de Martens e da Convenção sobre a Proibição ou Limitação do Uso de Certas Armas Convencionais[29]. O que nos remete, então, para esse plano de análise.

2. Armas Autónomas e Direito da Guerra

2.1 Nas AAs, a decisão militar é em certa medida 'tomada' pela própria máquina. Em conformidade com a programação que lhe foi dada, mas sem uma específica ordem humana na situação concreta[30]. Essas armas apresentam um elevado potencial de destruição, mas não residirá nisso a sua particularidade (basta pensar nas armas nucleares); antes no seu elevadíssimo grau de precisão. Embora sujeitas a erros próprios, estas armas operam livres daquela margem de erro humano que, mesmo para soldados com treino qualificado, é particularmente elevada em situações de grande tensão, que são, para eles, de vida ou morte[31] e em que, portanto, a força motivadora das normas jurídicas está muito mitigada.

É justamente essa assustadora precisão, no entanto, que torna plausível que a utilização de AAs assegure o 'cumprimento' da generalidade dos princípios e das regras do Direito da Guerra, como o princípio

---

[28] Cf. *v.g.* os §§ 15 e 32.

[29] Cf. *v.g.* os §§ 28 e 38.

[30] Cf., em grande medida para o que segue nesta secção, o já referido texto Miguel Lemos / Miguel João Costa, «Inteligência Artificial e Direito da Guerra», 91 ss.

[31] Ou, de forma distinta mas afinal concorrente para o mesmo sentido de fundo, em situações em que a decisão é inteiramente humana mas tomada remotamente (*v.g.*, através da operação de *drones* não autónomos): a tensão, aqui, não será comparável àquela que experiencia um soldado no terreno, mas há, em contrapartida, uma tendência para a desumanização do inimigo e, assim, para um menor refreamento no ataque: cf. Katalin Ligeti, "Artificial Intelligence and Criminal Justice", *Association Internationale de Droit Pénal*, Novembro de 2019 (aqui), 14-15, com ulteriores referências, falando a este respeito de "mentalidade PlayStation" e de "guerreiros de cubículo".

da distinção entre combatentes e civis e o princípio da proibição do sofrimento desnecessário. De forma que o movimento, entretanto surgido, no sentido de uma total proscrição destas armas não parece poder colher suporte expressivo no Direito da Guerra. Por não se tratar exactamente de um ramo de protecção de bens jurídicos, mas de minimização de danos – de "danos decorrentes de um fenómeno que não vê como aceitável, mas que presume ser inevitável e a que por isso procura oferecer-se como paliativo" –, o Direito da Guerra não só não se mostra especialmente adverso ao desenvolvimento e à utilização de AAs, como em certo sentido lhes dá mesmo algum suporte, pelas perspectivas que abrem de diminuição de fatalidades civis e de outros danos colaterais[32].

2.2 Também do ponto de vista da responsabilização penal por crimes de guerra as regras do Direito Internacional Humanitário, consuetudinário e codificado, proporcionam um enquadramento que pode ser visto como satisfatório (ainda que carecido de adaptações, o que, de todo o modo, não singulariza o âmbito militar de outros onde a utilização da IA enfrenta muito menor resistência, como o dos veículos autónomos ou o das intervenções médicas). Pense-se no crime de ataques indiscriminados: se o líder, alta patente, fabricante, comandante ou operador agiram de modo intencional e a AA foi utilizada com o objectivo de matar civis indiferenciadamente, todos eles (ou, pelo menos, algum ou alguns deles) terão cometido um crime de guerra; se havia razões para duvidar que certa AA tivesse a capacidade de diferenciar combatentes de civis e, apesar disso, não foram tomadas as precauções necessárias para evitar a sua utilização, serão também responsabilizáveis[33].

Há quem chame a atenção para a existência de uma "armadilha de responsabilidade"[34]: haverá aqui um conjunto de hipóteses, decorrentes de 'erros' do sistema de IA, em que nem a título de negligência se poderá estabelecer a responsabilidade de um ser humano. É essencial

---

[32] Cf. Miguel Lemos / Miguel João Costa, «Inteligência Artificial e Direito da Guerra», 118.

[33] Cf. Miguel Lemos / Miguel João Costa, «Inteligência Artificial e Direito da Guerra», 111 s.

[34] Cf. Afonso Seixas-Nunes, «Autonomous Weapons Systems and the Procedural Accountability Gap», *Brooklyn Journal of International Law* 46 (2021) 421-478 (disponível aqui).

dispensar atenção a este problema, mas não é claro que ele seja insuperável: desafios idênticos colocam-se, uma vez mais, e embora com variação nos seus contornos específicos, naqueles outros domínios onde a IA vem sendo aplicada sem resistência comparável.

2.3 Não pretende isto significar que não haja objecções sérias ao desenvolvimento e à utilização da IA no domínio militar – por exemplo: o risco de essas armas, justamente por não pressuporem o envio de forças humanas, diminuírem a contenção na decisão de intervir militarmente, causando maiores danos em termos absolutos (o que é um problema de *Jus ad Bellum* mais do que de *Jus in Bello*); o risco de uma corrida ao armamento, resultante na disseminação destas armas por actores (estaduais e não-estaduais) especialmente imprevisíveis; o risco de escalada de conflito, em virtude da enorme rapidez com que estas armas operam; o risco de discriminação algorítmica na identificação de alvos (*v.g.*, em razão da raça, do género, etc., risco também presente na utilização da IA em outros domínios, como o do policiamento preditivo e o da administração da justiça); o risco de estas armas desencadearem um processo de rejeição genérica da IA, comprometendo as virtudes que ela promete em outros sectores, como o da medicina; enfim, no limite, o próprio risco de a humanidade perder o controlo sobre a máquina[35].

Simplesmente, essa discussão não passa em primeira linha pelo Direito da Guerra – ou pelo Direito *tout court* –, mas por outros planos, como o da Ética, o da Filosofia da Ciência, o da Política ou o das Relações Internacionais. Daí poderão – deverão – emergir proposições merecedoras de positivação jurídica, dando já disto exemplo a Resolução do Parlamento Europeu acima referida, pese embora a sua reduzida força normativa. Mas o Direito da Guerra, em si, vê o desenvolvimento destas armas com uma certa naturalidade até, supondo que elas de facto apresentem a precisão e a capacidade de diferenciação que se anuncia e que, por essa via, proporcionem melhores resultados de um ponto de vista humanitário[36].

---

[35] Cf. Miguel Lemos / Miguel João Costa, «Inteligência Artificial e Direito da Guerra», 118 s.

[36] Cf. Miguel Lemos / Miguel João Costa, «Inteligência Artificial e Direito da Guerra», 119.

Este último é um ponto importante: o repúdio que vem sendo manifestado em relação às AAs baseia-se grandemente em dúvidas sobre a sua capacidade de alcançar os resultados visados pelo Direito da Guerra; esse argumento, porém, além de constituir uma questão mais científica do que jurídica – sobre que, portanto, nem posso pronunciar-me –, é notoriamente circular: se estas armas não tiverem essa capacidade, naturalmente que o Direito da Guerra se lhes oporá, pelo menos até certo ponto, e vice-versa[37]. A questão mais difícil é a de determinar se, *admitindo* que elas têm elas essa capacidade, não deverão ainda assim ser rejeitadas. E, de todo o modo, se é plausível, na crua realidade das coisas, que isso aconteça no mundo que temos. Se o não for – como julgo que não é –, a questão passará então a ser a de saber o que é que se pode fazer quanto a isso[38].

Conclusão

Em conclusão, creio se continuará inevitavelmente a desenvolver e a utilizar AAs, com menos constrangimentos do que noutros sectores onde a IA vem sendo aplicada. A Proposta de Regulamento da UE aponta claramente nessa direcção, ao excluir essa espécie de matérias do seu âmbito de aplicação e ao remeter o seu enquadramento jurídico, essencialmente, para o Direito Internacional Público, que não tem vocação para impor entraves significativos àquela utilização.

Referências bibliográficas:

Costa, Miguel João / Abrantes, António Manuel, "Os Desafios da Inteligência Artificial da Perspectiva Transnacional: A Jurisdição e a Cooperação Judiciária", in Anabela Miranda Rodrigues, ed., *A Inteligência Artificial no Direito Penal*, Coimbra: Almedina, 2020, 163-217.

---

[37] Cf. Miguel Lemos / Miguel João Costa, «Inteligência Artificial e Direito da Guerra», 100 s.

[38] Questão cujas implicações não ficam certamente apaziguadas, embora recebam ao menos um sinal de atenção, com o desenvolvimento de princípios de utilização responsável da IA no domínio militar como os recentemente adoptados pela NATO (aqui), §§ 7 ss.

Lemos, Miguel / Costa, Miguel João, "Inteligência Artificial e Direito da Guerra: Reflexões sobre as Armas Autónomas Mortíferas", in Anabela Miranda Rodrigues, coord., *A Inteligência Artificial no Direito Penal: Volume II*, Coimbra: Almedina, 2022, 91-123.

Ligeti, Katalin, "Artificial Intelligence and Criminal Justice", *Association Internationale de Droit Pénal*, Novembro de 2019, em https://www.penal.org/en/information

Llobet, Mariona, "Enemy Criminal Law (*Feindstrafrecht*)", in Pedro Caeiro / Sabine Gless / Valsamis Mitsilegas / Miguel João Costa / Janneke de Snaijer / Georgia Theodorakakou (eds.), *Elgar Encyclopaedia of Crime & Criminal Justice*, Cheltenham: Edward Elgar Publishing, 2023, em https://www.elgaronline.com/display/book/9781789902990/b-9781789902990.enemy.criminal.law.xml

Montag, Luca et al., "The Rise and Rise of Biometric Mass Surveillance in the EU: A Legal Analysis of Biometric Mass Surveillance Practices in Germany, the Netherlands, and Poland", EDRi (European Digital Rights)| Edinburgh International Justice Initiative, 2021, em https://edri.org/wp-content/uploads/2021/11/EDRI_RISE_REPORT.pdf

Rönnback, Ronja, "Challenges of Governing AI for Military Purposes and Spill-Over Effects of the AI Act", *The AI Alliance Blog*, 27 February 2023, em https://futurium.ec.europa.eu/en/european-ai-alliance/blog/challenges-governing-ai-military-purposes-and-spill-over-effects-ai-act

Seixas-Nunes, Afonso, "Autonomous Weapons Systems and the Procedural Accountability Gap", *Brooklyn Journal of International Law* 46 (2021) 421-478.

Outras fontes:

Amnesty International, "Forensic Methodology Report: How to catch NSO Group's Pegasus", 2021, em https://www.amnesty.org/en/documents/doc10/4487/2021/en/

European Center for Not-for-Profit Law, "Artificial Intelligence Act Amendments: Scope of the EU Artificial Intelligence Act (AIA): Military Purposes and National Security", 2022, em https://ecnl.org/sites/default/files/2022-03/ECNL%20Pagers%20on%20scope%20of%20AIA%20ECNL_FINAL.pdf

NATO, "Summary of the NATO Artificial Intelligence Strategy", 22 de Outubro de 2021, em https://www.nato.int/cps/en/natohq/official_texts_187617.htm

5

Inteligência artificial no âmbito da manutenção da ordem pública: considerações iniciais sob a ótica da proposta de Regulamento do Parlamento Europeu e do Conselho[1]

(https://doi.org/10.47907/DireitoemMudanca/2023/5)

*Alberto Raphael Ribeiro Magalhães**
*Ana Cristina Crestani***
*Luiza Tosta Cardoso Franco****

Resumo: O artigo compromete-se, ainda que superficialmente, a uma aproximação à análise da Proposta de Regulamento do Parlamento Europeu e do Conselho, que estabelece regras harmonizadas em matéria de IA, também no âmbito da manutenção da ordem pública. *Ab initio*, realiza-se uma abordagem acerca da inteligência artificial, perpassando, em seguida, pelo panorama protetor dos direitos fundamentais, com o apontamento da sua opacidade e considerável ameaça discriminatória,

[1] Este artigo foi escrito por pesquisadores do grupo de investigação "Inteligência Artificial e manutenção da Ordem Pública: impacto da proposta de regulamento de inteligência artificial no direito português" (projeto financiado no âmbito da iniciativa *Researchers' Camp*, do Instituto Jurídico), sob a tutoria da Professora Doutora Susana Aires de Sousa, sendo fruto de uma fase inicial desta investigação. O projeto, atualmente, encontra-se em desenvolvimento.

* Mestrando em Ciências Jurídico-Políticas, menção em Direito Constitucional, na Faculdade de Direito da Universidade de Coimbra; Orcid ID 0000-0002-5495-0461.

** Mestranda em Ciências Jurídico-Criminais na Faculdade de Direito da Universidade de Coimbra; Orcid ID 0009-0004-4528-2328.

*** Doutoranda em Direito na Universidade de Coimbra, no ramo de Direito Civil; Orcid ID 0000-0003-3982-1909.

particularmente relevantes nos sistemas de risco elevado para fins de manutenção da ordem pública, enumerados no item 6, do anexo III.

Palavras-chave: Inteligência Artificial; Ordem Pública; União Europeia; Proposta de Regulamento da Inteligência Artificial no âmbito da União Europeia.

1. Introdução

Na dialética de *estar* humano, o sujeito orgânico prossegue no *rio inovador heracliano*[2], desencadeando no sistema terrestre, por meio de distintas águas, uma nova esfera quando na corrida emancipatória de *ser* potência. A vida, uma das condições humanas sustentadas por *Hannah Arendt*, passa a englobar não mais os elementos estruturantes da hidro-atmo-lito-bio para admitir a «infosfera»[3] avistada por *Luciano Floridi*, pois, como seres condicionados, «tudo aquilo com o qual eles entram em contato torna-se imediatamente uma condição da sua existência»[4].

Assim, adota-se, ante esta nova circunstância ecossistêmica[5], o *global way of life* com a dinâmica *onlife*[6], ou seja, vive-se, ora direta ora

---

[2] Consoante o pensamento do filósofo pré-socrático "just as the river where I step is not the same, and is, so I am as I am not". (Heraclitus, of Ephesus, *Fragments: the collected wisdow of Heraclitus*, trad. Brooks Haxton, New York: Viking Penguin, 2011, 90).

[3] Nas palavras do filósofo italiano "The infosphere is the whole system of services and documents, encoded in any semiotic and physical media, whose contents include any sort of data, information and knowledge (...), with no limitations either in size, typology or logical structure. Hence it ranges from alphanumeric texts (i.e. texts including letters, numbers and diacritic symbols) and multimedia products to statistical data, from films and hypertexts to whole text-banks and collections of pictures, from mathematical formulae to sounds and videoclips". (Luciano Floridi, *Philosophy and computing: an introduction*, London and New York: Routledge, 2001, 8).

[4] Hannah Arendt, *A condição humana*, 10ª ed., Rio de Janeiro: Forense Universitária, 2007, 17.

[5] *Vide* Luciano Floridi, «Ethics after the information revolution», in Luciano Floridi, *The Cambridge handbook of information and computer ethics*, Cambridge: Cambridge University Press, 2010, 3-19, 8.

[6] Na atual evolução humana, cá estamos na posição de organismos vivos unidos numa experiência "onlife", já que a "(...) vida na infosfera, (...) não faz mais sentido perguntar se você está online ou offline, conectado ou não conectado". (Luciano Floridi, «Soft Ethics and the Governance od the Digitial», *Philosophy & Technology* 31/1 (mar. 2018), 1-8. [Consult. 17 abril 2023]. Disponível em: https://www.researchgate.net/publication/323248541_Soft_Ethics_and_the_Governance_of_the_Digital/link/5a895f23458515b8af92826f/download, 2).

indiretamente, conectados pelos múltiplos dispositivos e ferramentas digitais, a exemplo do telemóvel que carregamos em nossos bolsos, a *Alexa* mantida em nossa sala de estar ou pelos *smartwatches* nos pulsos, etecetera[7]. O Planeta Terra se tornou uma *World Wide Web*, já que, como num dilúvio técnico-científico-informacional, o mundo foi imerso e unificado nesta fluidez, recriando a *pangeia* de outrora sob a unidade da «information and communication technology»[8].

A *fumaça* utilizada por *modus* de comunicação rudimentar foi substituída pelas tecnologias da *digital revolution* desenvolvidas em laboratório, que entre erros e acertos, culminaram com a invenção de uma *criatura* que, aparentemente, supera o *criador* em termos de sapiência, malgrado — ainda — sob controlo do inventor[9]. Desse modo, a inteligência artificial (IA), antes objeto de temática ficcional, ultrapassou as telas no intuito de integrar as nossas vidas, assumindo um papel singular em diversos aspetos da vida indivíduo-social, como já se observa nos afazeres do quotidiano, na saúde com sistemas que antecipam diagnósticos de doenças e nos transportes com os automóveis autónomos. Em outras palavras, saiu-se do *imaginário* para uma «*re-ontologization of our environment and of ourselves*»[10].

Além do *locus* privado, alcança-se o terreno do público, mormente pela sua habilidade de *input* conhecimento, aprendendo pela experiência, e *output* soluções à vida social[11]. Por via de consequência, indubitável é que em alguns países o sistema de inteligência já é *eixo* da *engrenagem* voltada à manutenção da ordem pública, posto que, bem administrada, ela pode promover maior segurança à sociedade, com agilidade para resolução e prevenção de crimes. Contudo, num olhar crítico, jamais vendado, apesar dos benefícios latentes, fundamental é se debruçar sobre os riscos e as consequências da sua implementação

---

[7] Cf. Marcos Ehrhardt Jr. / Gabriela Buarque Pereira Silva, «Diretrizes éticas para a inteligência artificial confiável na união europeia», *Jurismat* 12 (nov. 2020), 305-337, 309;

[8] Floridi, *Philosophy and computing: an introduction*, 1.

[9] Manuel Lopes Rocha, «Nota prévia», in Manuel Lopes Rocha / Rui Soares Pereira, coord., *Inteligência artificial & direito*, reimp., Coimbra: Almedina, 2022, 5-9, 7-8.

[10] Luciano Floridi, «Ethics after the information revolution», 12.

[11] Alan Turing, *The essential Turing: seminal writings in computing, logic, philosophy, artificial intelligence, and artificial life plus the secrets of enigma*, editado por B. Jack Copeland, New York: Oxford University Press, 2004, 353.

nesta seara a fim de se obter uma inteligência artificial de confiança e *amiga* dos direitos fundamentais da espécie criadora.

Diante dessa realidade, o Parlamento Europeu conjuntamente com o Conselho da União Europeia, ao sopesar os *benefícios socioeconómicos* e *os riscos ou consequências negativas*, isto é, a «dupla face da IA»[12], apresentou uma sugestão regulamentar ao mercado interno de sistemas de IA, em 21 de abril de 2021, prezando sempre por um viés ético e valorativo em harmonia com a carta de direitos humanos[13], intitulado «Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de IA (Regulamento IA) e altera determinados atos legislativos da União»[14].

Para tanto, este trabalho realiza-se numa pesquisa qualitativa para análise da Proposta de Regulamento do Parlamento Europeu e do Conselho em matéria de IA, especificamente quanto ao enquadramento dos sistemas de inteligência no âmbito da manutenção da ordem pública, e pela compreensão da hermenêutica jurídica, tendo em vista a interpretação e integração das normas jurídicas[15].

2. Inteligência *extra*-humana

«*Qu'est ce qu'une intelligence artificielle?*»[16]. Tal seria o título da obra se *Sieyès* buscasse, no corrente período da humanidade em que

---

[12] Expressão utilizada por Susana Aires de Sousa para designar os dois lados (o bom e o mau) do sistema de IA. Para mais, Susana Aires de Sousa, «A IA no setor económico: uma reflexão entre o bom, o mau e o vilão», in Anabela Miranda Rodrigues, coord., *A inteligência artificial no direito penal* II, Coimbra: Almedina, 2022, 175-205, 181.

[13] União Europeia, *Carta dos Direitos Fundamentais da União Europeia*, 2016, [Consult. 17 abril 2023], Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:12016P/TXT&from=FR

[14] Parlamento Europeu/Conselho, *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021, [Consult. 17 abril 2023], disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:52021PC0206.

[15] Luís Poças, *Manual de Investigação em Direito: metodologia da preparação de teses e artigos jurídicos*. 2.ª ed, Coimbra: Almedina, 2022, 42.

[16] Paráfrase ao título *Qu'est-ce que le tiers état?*, publicado originalmente em 1789, do escritor-abade-político Emmanuel Joseph Sieyès. (Emmanuel-Joseph Sieyès, *¿Qué es el estado llano?: ensayo sobre los privilegios*; versão de José Rico Godoy, Madrid: Centro de Estudios Constitucionales, 1988).

o homem parece sair da posição de *sapiens* em direção à nova espécie *deus* do género *homos*, investigar sobre a temática e, decerto, em suas elucubrações, compartilharia da dificuldade de *Samir Merabet* em encontrar um uno conceito, já que «en l'état du droit positif, l'absence de désignations unique des technologies intelligentes et les carences des définitions utilisées concourent à entretenir l'incertitude».[17]

No entanto, ao autopsiar a *machine intelligence*[18] ou, agora habitualmente nomeada, *inteligência artificial*, constatará que integra ela dois fragmentos: inteligência e artificialidade. O primeiro retrata a capacidade de compreender, (re)conhecer, aprender e responder questões, «*criativas* e até *espontáneas*»[19], solucionando-as. O segundo é a máquina inventada à similaridade do *cérebro humano* para trabalhar através das *sinapses algorítmicas* decifrando códigos, dados ou sequências de etapas[20]. Desta maneira, seria ela um equipamento artificial de processamento de dados, por meio de algoritmos, cuja velocidade de trocas informacionais, viabiliza um *know-how* de informações nela depositadas e, com a plasticidade do *digital brain*, conquista horizontes além-pensar humanoide com independência e autossuficiência[21].

Desta forma, sublimando-se as dificuldades gerais, especialmente à pretensão do Regulamento em repercutir em 27 (vinte e sete) ordens jurídicas soberanas e independentes da União Europeia, a sua proposta definiu, no seu artigo 3.º, o sistema de inteligência artificial como sendo «um programa informático desenvolvido com uma ou várias das técnicas e abordagens enumeradas no anexo I, capaz de, tendo em vista um determinado conjunto de objetivos definidos por seres humanos,

---

[17] Samir Merabet, *Vers un droit de l'intelligence artificielle,* Paris: Dalloz, 2020, 49. Nessa mesma lógica: Ehrhardt Jr./Silva, «Diretrizes éticas para a inteligência artificial confiável na união europeia», 307.

[18] Termo original cunhado por Alan M. Turing para se reportar o que é conhecido nomeadamente como inteligência artificial, ainda em uso na Grã-Bretanha. *Ver* Turing, *The essential Turing: seminal writings in computing, logic, philosophy, artificial intelligence, and artificial life plus the secrets of enigma*, 353.

[19] Susana Aires de Sousa, «"Não fui eu, foi a máquina": teoria do crime, responsabilidade e inteligência artificial», in Anabela Miranda Rodrigues, coord., *A inteligência artificial no direito penal*, Coimbra: Almedina, 2020, 59-93, 75.

[20] James H. Fetzer, *Artificial Intelligence: Its scope and limits*. Kluwer Academic Publishers. 1990, 3-4.

[21] Jerry Kaplan, *Artificial Intelligence: what everyone needs to know*. Oxford: Oxford University Press, 2016, 4.

criar resultados, tais como conteúdos, previsões, recomendações ou decisões, que influenciam os ambientes com os quais interage»[22].

3. *Proposta de Regulamento* da IA

Em sua exposição de motivos, a *proposição regulamentar europeia* explicita a inteligência artificial consoante «uma família de tecnologias em rápida evolução capaz de oferecer um vasto conjunto de benefícios económicos e sociais a todo o leque de indústrias e atividades sociais»[23], sem olvidar da potencialidade dos riscos inerentes à sociedade. Deixa-se, nitidamente, percetível o desassossego de sua opacidade, complexidade, comportamento autónomo e com o enviesamento que podem advir dos sistemas de IA[24].

Neste diapasão, assume-se como um dos seus objetivos o afiançar de mecanismos seguros e garantidores dos direitos e valores fundamentais, pois, malgrado o belo *cântico sinero artificial*, que nos encanta e hipnotiza pela sua possibilidade *absoluta* de perfeição, cuidado com o desconhecido e suas afetações aos direitos citadinos não caracteriza um excesso dispensável, mas sim imprescindível em decorrência aos preconceitos e comportamento quase humanos já constatados.

Por isso, como *essência do funcionamento*, delimita-se um campo de atuação do sistema de IA em que o território dos direitos fundamentais não pode ser *esbulhado*, sendo, numa inversão do paradigma maquiavélico, os *meios* (requisitos sistémicos baseados numa análise de risco)

---

[22] Parlamento Europeu/Conselho, *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021, [Consult. 17 abril 2023], disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:52021PC0206.

[23] Parlamento Europeu/Conselho, *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021, [Consult. 17 abril 2023], disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:52021PC0206, 1.

[24] Parlamento Europeu/Conselho, *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021, [Consult. 17 abril 2023], disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:52021PC0206, 1.

como justificativas dos *fins* (proteção da ordem pública e promoção dos direitos descritos na Carta de Direitos Fundamentais europeus) que se propõe com a implementação e atuação da *inteligência artificial* na União Europeia[25].

3.1 Os «sistemas de IA»: *«um novo mundo ao homem, um «upgrade» à humanidade»*[26]

Criar um quadro regulamentar é preciso, proclamaria (as) *Pessoa*(s) – o Fernando e/ou todos nós – ao avistar o *artificial being* deslocando neste solo ainda não delimitado, sem limites de *como* e se *pode agir*, a fim de proporcionar uma convivência pacífica e harmoniosa com os fundamentais direitos, ora individuais ora coletivos, e, consequentemente, a segurança pública. Mas, como qualquer ato administrativo, ainda que seja a nível europeu, a *medida* adotada deve guardar proporção com o escopo pretendido, evitando-se excesso da autoridade. *Übermassverbot,* ou princípio da proibição do excesso, foi «europeizado», conforme *Canotilho*, exercendo uma *função de controlo* consentâneo ao equilíbrio entre a *prossecução finalística* e os *direitos e interesses em questão*[27].

Sob este *ponto do direito europeu*, confiou-se ao *Livro Branco* o detalhamento das características essenciais a ser imposta na qualidade de requisitos mandamentais aos sistemas de IA (dados de treino, conservação de registo de dados, prestação de informações, robustez e exatidão, supervisão humana e requisitos específicos para pontuais aplicações de IA) a fim de se evitar desencontros com os direitos dos cidadãos europeus[28]. Assim, a União Europeia, em seus primeiros

---

[25] Parlamento Europeu/Conselho, *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021, [Consult. 17 abril 2023], disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:52021PC0206, 3-7.

[26] Paráfrase ao dizer de *Neil Armstrong* ao pisar na Lua em 20 de julho de 1959.

[27] J. J. Gomes Canotilho, *Direito constitucional e teoria da constituição*, 7ª ed., 11ª reimp., Coimbra: Almedina, 2003, 268.

[28] Comissão Europeia, *Livro branco sobre a Inteligência Artificial - uma abordagem europeia virada para a excelência e a confiança*, 2020, [Consult. 17 abril 2023], Disponível em: https://op.europa.eu/pt/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1, p.20.

passos ao equilíbrio na regulamentação, demonstra-se consciente dos *risco*s aos direitos e valores fundamentais, como também cônscio da necessidade de favorecer os investimentos e inovações no domínio da IA[29]. Dessarte, após estudos e pesquisas realizadas por um grupo de *experts* independentes (GPAN IA[30]), elegeu-se uma classificação dos sistemas de risco de acordo com os parâmetros da probabilidade e da proporcionalidade da violação de direitos fundamentais e da segurança jurídica, bem como pela funcionalidade, finalidade e modalidade do uso[31]. São eles: (i) *risco inaceitável*; (ii) *risco elevado*; e (iii) *risco baixo ou mínimo*[32].

Nesse sentido, «a utilização de sistemas de alto risco ou é proibida pela Proposta da Comissão ou deverá cumprir requisitos apertados consagrados na Proposta»[33], uma vez que, ao se observar a prescrição do artigo 5°, n. 1, alínea *d*, da Proposta de Regulamento, proíbe-se o «uso de sistemas de identificação biométrica à distância em «tempo real» em espaços acessíveis ao público para efeitos de manutenção da ordem pública» – embora limitadas exceções são erigidas, as quais devem guardar semelhança aos requisitos da *Proposta*[34]. Já os sistemas

---

[29] *Ver* item 5.2 da *Exposição de Motivos* e o art. 5.º do *Título II* da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021.

[30] Exposição de Motivos da Proposta de Regulamento, páginas 9 e seguintes. Cf. Parlamento Europeu; Conselho. *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021, [Consult. 17 abril 2023], Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:52021PC0206.

[31] Equitativamente, no *Livro Branco* os sistemas de IA para serem reputados de risco elevado devem preencher, cumulativamente, estes dois requisitos.

[32] *Ver* item 5.2 da *Exposição de Motivos* e o art. 5.º do *Título II* da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021.

[33] Rui Soares Pereira, «Sobre o uso de sistemas de identificação biométrica (e de tecnologias de reconhecimento facial) para fins de segurança pública e de aplicação coercitiva da lei: reflexões a propósito da proposta de regulamento europeu sobre a inteligência artificial», *Revista da Faculdade de Direito da Universidade de Lisboa* 63/1-2 (2022) 839-865, 860-861.

[34] *Ver* art. 5.°, n.° 2 a 4, da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021.

classificados como de risco mínimo, sujeitam eles aos pressupostos de *transparência e informação*, especialmente quando se versa de um *sistema biométrico de categorização*[35], devendo estar os utilizadores cientificados de estão a interagir com um sistema de IA e que que suas emoções e/ou suas características físicas-comportamentais estão a ser reconhecidas pelo mesmo, salvo seja ela manuseada, sob autorização legal, para detetar, prevenir e investigar infrações penais[36].

A normativa da *Proposta* arquiteta diretrizes claras para a utilização dos sistemas de IA para fins de manutenção da ordem pública, abrangendo ordens, por exemplo, relacionadas ao reconhecimento facial. Entretanto, reforça-se que ela não se aplica aos sistemas de IA desenvolvidos ou usados exclusivamente para fins militares. Neste caso, a competência regulamentar está adstrita ao Título V do Tratado da União Europeia (TUE), isto é, as *disposições gerais relativas à ação externa da União e disposições específicas relativas à política externa e de segurança comum*, e não o regulamento – art. 2.°, n.° 3. Sendo assim, se faz necessário averiguar para qual finalidade será ela utilizada, verificando-se se deverá seguir as regras previstas na Proposta ou, então, do Direito Internacional Público, especificamente no direito da guerra.

Excluindo-se a excecionalidade militar, com a *propositura* de requisitos e obrigações a serem cumpridas, certas tipificações, como as identificadas sendo de *risco inaceitável*[37], foram vedados pelo *high level* de periculosidade de ultraje dos direitos fundamentais, enquanto, por outro lado, apesar de demonstrar um *risco elevado*[38]de violação, carece ser sopesado se o setor e a utilização do sistema encerram riscos

---

35 *Ver* artigo 52.° da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021.

36 Rui Soares Pereira, «Sobre o uso de sistemas de identificação biométrica (e de tecnologias de reconhecimento facial) para fins de segurança pública e de aplicação coercitiva da lei: reflexões a propósito da proposta de regulamento europeu sobre a inteligência artificial», 860-861.

37 *Ver* item 5.2.2 da *Exposição de Motivos* da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021.

38 *Ver* item 5.2.3 da *Exposição de Motivos* da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021.

significativos, nomeadamente aqueles voltados à proteção da segurança, aos direitos dos consumidores e dos direitos fundamentais. E, por fim, não menos importante, porém mais seguros aos indivíduos, são os sistemas de *risco baixo/mínimo*, os quais são aqueles que possuem menor, ou, quase nenhuma, probabilidade de ferir os *valores* europeus.

Logo, com os *grandes poderes* dispensados pela Inteligência Artificial, *responsabilidade* e *controlo* são as chaves para manutenção do *status quo* direitos e valores fundamentais ao grande público europeu sob a compreensão de que esta tecnologia representa um *novo mundo ao homem, um «upgrade» à humanidade*.

3.2 Inteligência artificial: o *novo lobo* do homem?

O homem, singular, ao se descobrir como o *lobo* do seu igual, num consenso de todo o *cœtus*, criou o primeiro *ente artificial*, o *Leviathan*, para proporcionar à turba amedrontada, tanto pela incerteza jurídica quanto pela submissão ao mais forte, três elementos finalísticos desta abstração composta por um território, povo e poder único e concentrado: *segurança, justiça e bem-estar*[39] numa escritura.

Da *teoria orgânica* à *teoria da integração*, «o próprio Estado constituía, (...) um todo, animado de vida própria»[40], cujo processo de formação adveio de uma vontade geral. Um *ser* abstrato, dotado de personalidade, que ultrapassa os indivíduos, porém cerceado por um programa de normas. Para, em seguida, caracterizar-se como um *todo composto por uma rede unitária integrada* e, por sua vez, realocar a condição de existir na manutenção diária do querer dos seus administrados por meio do ordenamento jurídico e sua respetiva constituição política[41]. Independentemente da teoria, o *Estado* é uma "tecnologia política de equilíbrio político-social"[42], no qual reúne em si o monopólio

---

[39] Para António Pedro Ribeiro dos Santos é o entrelaçar desses três pontos que correspondem, na grande parte dos autores voltados à ciência política, a finalidade do Estado. *Ver* António Pedro Ribeiro dos Santos, *O estado e a ordem pública: as instituições militares portuguesas*, Lisboa: Instituto Superior de Ciências Sociais e Políticas, 1999, 11.

[40] Reinhold Zippelius, *Teoria geral do estado*, trad. de António Cabral de Moncada, 2ª ed., Lisboa: Fundação Calouste Gulbenkian, 1995, 23.

[41] Cfr. Reinhold Zippelius, *Teoria geral do estado*, 34-35.

[42] Canotilho, *Direito constitucional e teoria da constituição*, 89.

legiferante e coercitivo para a defesa da – *ancestralidade napoleônica*[43] – *ordem pública*, cujo conceituação seja difícil destrinçar, seus objetivos, por outro lado, são cognoscíveis, pois, sincronicamente, deve-se salvaguardar a paz entre os sócios e respeitar as liberdades dos mesmos, sem qualquer ação contrária aos fundamentais direitos[44].

Ao mirar tal inconveniente, a *proposta* de regulamentação deste *ente artificial*, que, em caso de aprovação, reverberará em outras ordens internas por meio de regras de direito internacional, não pelo *consensus omnium*, mas pela ação de *poucos*, definiu, em seu art. 3º, n.º 41, a *ordem pública* como «às atividades realizadas por autoridades policiais para efeitos de prevenção, investigação, deteção ou repressão de infrações penais, ou execução de sanções penais, incluindo a proteção contra ameaças à segurança pública e a prevenção das mesmas»[45].

Apreensivos pela assimetria de poder que se revela às *autoridades policiais*[46] ante o domínio da *nova entidade*, cujo condão oportuniza «à vigilância, detenção ou privação da liberdade de uma pessoa singular, bem como ter outros impactos adversos nos direitos fundamentais garantidos pela Carta»[47], reclama-se um *equilíbrio* com preceitos estaduais valorativos, diretivos e principiológicos estabelecidos pela *primazia do coletivo*[48].

---

[43] De acordo com Julio O. de Roa, embora a noção de *ordem pública* tivesse possa ser percebida no *Digesto*, foi no *Code Napoléon* que principiou categoricamente o conceito em um corpo legislativo. Para mais, *ver* Julio O. de. Roa, *Del orden publico em derecho positivo*, Buenos Aires: Librero Editor, 1926, 1.

[44] Cfr. Jacob Dolinger, *A evolução da ordem pública no direito internacional privado*, Rio de Janeiro: Almedina, 1979, 13.

[45] Parlamento Europeu/Conselho, *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021, [Consult. 17 abril 2023], Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:52021PC0206.

[46] Consoante o *Regulamento*, leia-se toda autoridade pública ou organismo/entidade designado pelo Estado-membro com competência para exercer as referidas funções (art. 3.º, n.º 40).

[47] Considerando n. 38 da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021.

[48] Cfr. Tomás Prieto Álvarez, *La dignidade de la persona: núcleo de la moralidad y el orden públicos, límite al ejercicio de libertades públicas*, Navarra: Editorial Aranzadi, 2005, 54.

Viver é *arriscar*-se. Com isso, entre *riscos*, percorre(u) a humanidade o caminho do ininterrupto progresso, no qual, ao se submeter no *duelo* diacrónico, busc(a/ou)-se o *controlo* do *desconhecido* até *superar* os seus perigos advindos. Não diferente, o sistema de IA, que – certamente – provoca(rá) uma remodelação da vida social e do próprio sujeito[49]–, traz consigo um compêndio de *riscos*, cuja *intensidade* e *tipo* se materializam de acordo com o seu uso e destinação[50]. Ciente de tal, o *Direito*, regramento voltado à *tríplice finalidade de justiça, segurança e bem-estar*, por estar a enfrentar um *novo lobo – incerto – dos homens*, precisa guarnecê-los de «medidas, obrigações e proibições, ao abrigo de um *princípio da precaução* – deveres necessariamente atualizáveis e passíveis de revisão à medida que se tornam mais conhecidos»[51].

Com especial apreço, *a proposta regulamentar* destinou no anexo III a apresentação pormenorizada do rol dos sistemas elencados de *risco elevado*, cuja materialização da ameaça ora já ocorrido ora se encontra suscetível de ocorrer em um futuro próximo, o qual é referenciado no art. 6.º, n.º 2, dispondo ser aqueles voltados à «identificação biométrica e categorização de pessoas singulares[52]», «gestão e funcionamento de infraestrutura», «educação e formação profissional», «emprego, gestão de trabalhadores e acesso ao emprego por conta própria», «acesso a serviços privados e a serviços e prestações públicas essenciais, bem como o usufruto dos mesmos», «gestão da migração, do asilo e do controlo das fronteiras», «administração da justiça e processos democráticos» e «manutenção da ordem pública»[53].

---

[49] Cf. Tomás Prieto Álvarez, *La dignidade de la persona*, 193.

[50] Cf. Susana Aires de Sousa, «A IA no setor económico: uma reflexão entre o bom, o mau e o vilão», 191; Idem, «"Não fui eu, foi a máquina": teoria do crime, responsabilidade e inteligência artificial», 85-86.

[51] Susana Aires de Sousa, «A IA no setor económico: uma reflexão entre o bom, o mau e o vilão», 193.

[52] Cf. Rui Soares Pereira, «Sobre o uso de sistemas de identificação biométrica (e de tecnologias de reconhecimento facial) para fins de segurança pública e de aplicação coercitiva da lei: reflexões a propósito da proposta de regulamento europeu sobre a inteligência artificial», 841-842, a Proposta de Regulamento parece trazer uma abordagem bastante restritiva aos sistemas de identificação biométrica quando comparada às abordagens encontradas em outros países.

[53] Parlamento Europeu/Conselho, *Anexos da Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021, [Consult. 17 abril 2023], Disponível em: https://eur-lex.

No que concerne *ao último* – porém não mais importante pois corresponde ele ao objeto deste trabalho –, evidencia-se que o manuseamento de sistemas de IA incluído no domínio da manutenção da ordem pública possui alta potencialidade em causar grave ameaça quando utilizados à deliberação sobre temáticas que tocam os direitos fundamentais das pessoas, especialmente os direitos à liberdade e à privacidade, violando, por via de consequência, princípios, tanto de ordem constitucional quanto processual penal.

Por isso, pôs-se uma listagem *atualizável*, já que poderá ser posteriormente incrementada pela Comissão[54], conciliando o poder delegado conferido pelo Comité[55], onde se indica explicitamente, conforme se verifica no *ponto 6 do Anexo III*, os sistemas de IA, pertinentes à manutenção da ordem pública, considerados como *risco elevado*. São eles aqueles utilizados por autoridades policiais (*i*) «em avaliações individuais de riscos relativamente a pessoas singulares, a fim de determinar o risco de uma pessoa singular cometer infrações ou voltar a cometer infrações ou o risco para potenciais vítimas de infrações penais»[56]; (*ii*) «como polígrafos e instrumentos similares ou para detetar o estado emocional de uma pessoa singular»[57]; (*iii*) «para detetar falsificações

---

europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0004.02/DOC_2&format=PDF, 5-6

[54] Cfr. Art. 7.º da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União,* 2021.

[55] Cfr. Art. 73 e seguintes da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União,* 2021.

[56] Alínea *a* do ponto 6 do Anexo III da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021.

[57] A alínea *b* do ponto 6 do Anexo III aborda o *reconhecimento-identificação/aferição de emoções*. A *proposta regulamentar* dispõe acerca da necessidade da ciência do usuário acerca da execução deste tipo de sistema, salvo quando a categorização biométrica seja legalmente autorizada para detetar, prevenir e investigar infrações penais. Dentre tais sistemas, há o *polígrafo*, cujo método de *detecção da verdade* (ou da mentira) é discutível em virtude da possível violação aos princípios do contraditório, da imediação e proibição da autoincriminação. *Vide* Susana Aires de Sousa, *Neurociências e direito penal: em busca da verdade perdida (na mente)*, Coimbra: Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2017, [Consult. 17 abril 2023], disponível em: https://www.uc.pt/site/assets/files/435430/direitonumahora2_neurocie_ncias_e_direito_penal.pdf., 10-11.

profundas referidas no artigo 52.º, n.º 3»[58]; (*iv*) «para avaliar a fiabilidade dos elementos de prova no decurso da investigação ou repressão de infrações penais»[59]; (*v*) «para prever a ocorrência ou a recorrência de uma infração penal real ou potencial com base na definição de perfis de pessoas singulares, na aceção do artigo 3.º, ponto 4, da Diretiva (UE) 2016/680, ou para avaliar os traços de personalidade e as características ou os comportamento criminal passado de pessoas singulares ou grupos»[60]; (*vi*) «para definir o perfil de pessoas singulares, na aceção do artigo 3.º, ponto 4, da Diretiva (UE) 2016/680, no decurso da deteção, investigação ou repressão de infrações penais»[61]; (*vii*) «para serem utilizados no estudo analítico de crimes relativos a pessoas singulares, permitindo às autoridades policiais pesquisar grandes conjuntos de dados complexos, relacionados ou não relacionados, disponíveis em diferentes fontes de dados ou em diferentes formatos de dados, no intuito de identificar padrões desconhecidos ou descobrir relações escondidas nos dados»[62].

---

[58] Alínea *c* do ponto 6 do Anexo III da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*. 2021.

[59] Alínea *d* do ponto 6 do Anexo III da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*. 2021.

[60] Alínea *e* do ponto 6 do Anexo III da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*. 2021

[61] Alínea *f* do ponto 6 do Anexo III da *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*. 2021.

[62] A alínea *g* do tópico 6 do Anexo III corresponde ao denominado *sistema de policiamento preditivo*, cuja ferramenta, através de dados cronológicos, cria uma previsão espaço-tempo sobre as zonas de criminalidade e/ou prática de crime através de padrões e tendências, viabilizando estratégia assertiva na destinação de força policial para desarticular atividade criminosa. (Cfr. Jerry Ratcliffe, «What is the future… of predictive policing?», *Translational Criminology*, spring 2014, 4–5, 4). Em igual sentido, Christoph Burchard, «Artificial Intelligence as the End of Criminal Law? On the Algorithmic Transformation of Society» in Maria João Antunes / Susana Aires de Sousa, *Artificial Intelligence in the economic sector prevention and responsibility*, Coimbra: Instituto Jurídico Faculdade de Direito da Universidade de Coimbra, 2021,

Carece entender que o *tudo não valerá a pena se o cuidado for pequeno*[63], porque os sistemas, ao serem submetidos à técnica *machine learning*[64], na qual os dados processados (algoritmos) integram o seu parâmetro *funcional decisório*[65], realizando por si escolhas ora decorrentes da ação de programadores – *indivíduos que possuem suas preconceções que, ou, sem cautela, ainda que indiretamente*[66], *podem lançar à*

165-205, 175-184, traz-nos uma listagem das promessas da inteligência artificial: (i) eficácia e eficiência na inibição do crime, principalmente baseado no policiamento preditivo; (ii) objetividade, neutralidade e coerência na aplicação do Direito Penal. Embora o termo *policiamento preditivo* possa ser utilizado para abranger três categorias, quais sejam: (i) previsão dos perpetradores, (ii) das vítimas, e (iii) quando/onde incide maior risco de ocorrência criminosa, seu uso corriqueiro tende a fazer alusão a última categoria. Para mais: Wim Hardyns / Anneleen Rummens, «Predictive Policing as a New Tool for Law Enforcement? Recent Developments and Challenges», *European Journal on Criminal Policy and Research* 24 (2018) 201-218, 203); Para mais acerca do tema policiamento preditivo ver Jennifer Bachner, *Predictive policing: preventing crime with data and analytics*, Washington: IBM Center for the business of government, 2013, [Consult. 17 abril 2023]. Disponível em: http://www. businessofgovernment. org/sites/default/files/Predictive%20Policing.pdf. Acesso em 17 abril 2023).

63 Paráfrase ao poema de Fernando Pessoa «Mar Português» retirado de http:// arquivopessoa.net/textos/2405.

64 Mafalda Miranda Barbosa afirma que "a *machine learning* surge, assim, como um campo das ciências da computação que estuda os programas de computadores capazes de aprender com base na experiência e, assim, capazes de desenvolver as suas próprias potencialidades ao logo dos tempos. Trata-se, como os autores sublinham, de uma aprendizagem em termos funcionais: a modificação do comportamento do algoritmo dá-se como forma de desenvolvimento do seu desempenho na realização de certa tarefa e opera através da experiência" (Mafalda Miranda Barbosa, *Inteligência Artificial: entre a utopia e a distopia, alguns problemas jurídicos*, Coimbra: Gestlegal, 2021, 209). Sobre isso, *Criado* e *Such* afirmam que «Direct Discrimination (also known as Disparate Treatment) considers the situations in which an individual is treated differently because of their membership to a particular social group. This ultimately means that different social groups are being treated differently, with some of them effectively being disadvantaged by these differences in treatment» (Natalia Criado; Jose M. Such, «Digital discrimination», in Karen Yeung /Martin Lodge, *Algorithmic regulation*, 2019, 82-97, 83).

65 Cf. Susana Aires de Sousa, «"Não fui eu, foi a máquina": teoria do crime, responsabilidade e inteligência artificial», 60;

66 Segundo *Criado* e *Such*, «Indirect Discrimination (also known as Disparate Impact) considers the situations in which an apparently neutral act has a disproportionate negative effect on the members of a particular social group. This is considered discrimination even if there is no intention to discriminate that particular group or if there is not any unconscious prejudice motivating the discriminatory act» (Criado; Such «Digital discrimination», 83).

máquina –, ora desenvolvidas autonomamente pelos mesmos[67], podem provocar uma *digital discrimination*[68] e, por sua vez, uma violação de *bens juridicamente fundamentais*.

Desta forma, deve ser ponderada, quando eles são empregues em avaliações *individuais ou coletivas*, a potencialidade de consecução de fins *transgressores*, tanto pela *vigilância omnipresente*, pelo enviesamento discriminatório, quanto pela abstenção da transparência nos percursos "neurais" da IA até o seu *decisum*, haja vista a *possibilidade* de apresentar uma solução «unfairly, unethically or just differently based on their personal data such as income, education, gender, age, ethnicity, religion»[69].

Uma das maiores preocupações acarretadas pelos sistemas de inteligência artificial, sobretudo nas operações de manutenção da ordem pública, quando se concatena aos sistemas de utilização – ou que virão a ser utilizados – pelas autoridades policiais, é a designada, supracitada, porém não explicitada, *vigilância omnipresente*, em melhores palavras, vigilância estatal ou massiva aos utilizadores[70] de um *Big Brother* à la George Orwell, em que todos os passos objetivos e subjetivos, inclusive o mais particular, de uma pessoa estaria sendo *acompanhado* por uma autoridade. Receia-se, por via de consequência, uma invasão ao espaço privado, cuja redução não se pode alcançar ao cenário quimérico, impondo-se a verificação do potencial risco e impacto aos direitos fundamentais, como à liberdade e à segurança individual[71], ante, em

---

[67] Sobre tal, Susana Aires de Sousa, «"Não fui eu, foi a máquina": teoria do crime, responsabilidade e inteligência artificial». 59-93, onde desenvolve uma análise minuciosa sobre as hipóteses em que a máquina elabora, com independência, informação sequer programada e previsível pelo homem.

[68] Para aceder à informação mais completa, *ver* Criado / Such «Digital discrimination», 84.

[69] Criado / Such «Digital discrimination», 82.

[70] Rui Soares Pereira, «Sobre o uso de sistemas de identificação biométrica (e de tecnologias de reconhecimento facial) para fins de segurança pública e de aplicação coercitiva da lei: reflexões a propósito da proposta de regulamento europeu sobre a inteligência artificial», 841 e 863.

[71] Referindo-se aos sistemas biométricos e de reconhecimento facial, Pereira aponta que «existe uma forte possibilidade dos sistemas de reconhecimento facial serem usados para além do fim inicialmente autorizado e controlado e, consequentemente, este facto poderá: (i) colocar em risco a possibilidade de movimentação no espaço público de forma anónima; (ii) determinar um conformismo prejudicial ao livre-arbítrio; (iii) afetar as liberdades religiosas e os direitos das crianças; (iv) interferir com a liberdade de opinião e expressão da pessoa e ter um efeito negativo no direito de reunião e de associação; (v) ter um forte impacto no comportamento social e psicológico dos cidadãos; e (vi) sublinhar questões éticas importantes'». (Pereira,

termos legais, a deficiência normativa[72], e, numa perspetiva técnica, os «(…) graus de fiabilidade e precisão muito variados e um impacto na proteção dos direitos fundamentais e na dinâmica dos sistemas de justiça criminal»[73] do sistema de IA.

Destarte, a *caixa preta da* IA, denominada *black box problem* – uma espécie de *caixa de Pandora* do tempo corrente por carregar, apesar da *esperança* no porvir, eventuais *males* à humanidade – revela, até este momento, a opacidade do percurso do tratamento algoritmo de dados ao produto final. O caminho entre o *input* até o *output*, desconsiderando a precisão da informação estruturada, não é acompanhado de um *mapa*, mas de uma *confiança cega* pela ausência de transparência e explicação das decisões tomadas *algoritmicamente*[74]. O homem, *curioso*, abriu-a, sem um regramento prévio, deixando escapar os perigos *aos direitos fundamentais*, em especial à dignidade, igualdade sob o manto da não discriminação, proteção de dados e outros, que, ao contato do mesmo, contaminou-se com os pré-conceitos, enviesando-se numa orientação hostil.

Assim, para se evitar um alto custo ao indivíduo – ante a dificuldade de se fazer materializar um *direito à explicação*[75] *e à oposição* da

---

«Sobre o uso de sistemas de identificação biométrica (e de tecnologias de reconhecimento facial) para fins de segurança pública e de aplicação coercitiva da lei: reflexões a propósito da proposta de regulamento europeu sobre a inteligência artificial», 847).

72 Rui Soares Pereira, «Sobre o uso de sistemas de identificação biométrica (e de tecnologias de reconhecimento facial) para fins de segurança pública e de aplicação coercitiva da lei: reflexões a propósito da proposta de regulamento europeu sobre a inteligência artificial», 841 e 863.

73 Considerando M, da Resolução do Parlamento Europeu, de 06 de outubro de 2021. Parlamento Europeu/Conselho. *Resolução do Parlamento Europeu sobre a inteligência artificial no direito penal e a sua utilização pelas autoridades policiais e judiciárias em casos penais*, 2021, [Consult. 17 abril 2023], Disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2021-10-06_PT.html.

74 Susana Aires de Sousa, «"Não fui eu, foi a máquina": teoria do crime, responsabilidade e inteligência artificial», 66-67.

75 Cf. Susana Aires de Sousa, «"Não fui eu, foi a máquina": teoria do crime, responsabilidade e inteligência artificial», 67, expõe que, consoante a falta de transparência, «no plano jurídico, (...) emergi[u] em alguma literatura a apologia de um reconhecimento do "direito à explicação" no contexto de decisões automáticas tomadas por algoritmos». Assim, o *Regulamento geral sobre proteção de dados da União Europeia* de n. 679/2016 assegurou tais direitos em seus artigos 15.º, 21.º e 22.º (https://www.pgdlisboa.pt/leis/lei_mostra_articulado.php?nid=2961&tabela=leis). (União Europeia, *Regulamento geral sobre a proteção de dados da União Europeia* n.

decisão do sistema de IA, especialmente quando se dirige aos perfis individuais[76], embora espera-se uma mesma medida no que tange à uma afetação coletiva –, correta foi a medida adotada pela *União Europeia* no *Capítulo 2* do *Título III* da *Proposta*, que, ao invés de simplesmente rotulá-las como se inaceitáveis fosse, cerrando os olhos ao inevitável amanhã digital, estipulou como *conditio sine qua non* as incumbências seguintes para a implementação e utilização dos *sistemas de IA* em matéria de ordem pública: (i) *criação*, *implementação*, *documentação* e *mantimento de um sistema de gestão de riscos* (art. 8.º); (ii) *rastreabilidade* (art. 12.º); (iii) *transparência* e *prestação de informações aos utilizadores* (art. 13.º); (iv) *supervisão humana* (art. 14.º); (v) *exatidão, solidez e cibersegurança* (art. 15.º)[77]. Equitativamente, o compromisso dos mesmos em assegurar os direitos prescritos na *Carta dos Direitos Fundamentais* e no *Regulamento geral sobre a proteção de dados* da UE e não discriminar, direta ou indiretamente[78], por meio dos *neurónios* (*artificiais)* de um *res*[79] eletrónico.

679/2016, [Consult. 17 abril 2023], Disponível em https://www.pgdlisboa.pt/leis/lei_mostra_articulado.php?nid=2961&tabela=leis).

[76] Art. 22.º do Regulamento geral sobre a proteção de dados. *Vide* União Europeia, *Regulamento geral sobre a proteção de dados da União Europeia* n. 679/2016, [Consult. 17 abril 2023], Disponível em https://www.pgdlisboa.pt/leis/lei_mostra_articulado.php?nid=2961&tabela=leis.

[77] Embora não seja intenção nossa e, até mesmo, nosso objetivo explicitar cada *requisito*, importa-se entender que (i) a *transparência* exige que «os utilizadores devem ser capazes de interpretar o resultado do sistema e utilizá-lo de forma adequada», ou seja, «devem ser acompanhados de documentação pertinente e instruções de utilização e incluir informações concisas e claras» (art. 13.°); (ii) a *supervisão humana* impõe ao fornecedor de modo a «garantir que o sistema integre restrições operacionais que não possam ser neutralizadas pelo próprio sistema e que respondam ao operador humano e que as pessoas singulares a quem foi atribuída a supervisão humana tenham as competências, a formação e a autoridade necessárias para desempenhar essa função» (art. 14°); *solidez* significa que o sistema deve ser resistente a riscos, falhas, incoerências e erros, bem como ser resistente à «ações maliciosas suscetíveis de pôr em causa a segurança do sistema de IA» (art. 15.º, n°3); e a *cibersegurança* visa garantir que «os sistemas de IA sejam resistentes às ações de terceiros mal-intencionados que tentam explorar as vulnerabilidades dos sistemas» (art. 15, n.° 4).

[78] Parlamento Europeu/Conselho, *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021, [Consult. 17 abril 2023], Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:52021PC0206, 11.

[79] Ainda que se busque humanizar e sustentar uma *personalidade jurídica eletrônic*a à máquina, esta ainda, a priori, é um produto, uma coisa, isto é, pode ser – e é

Diante disso, com a tocha da *esperança* acesa, nessa busca por construir um *ecossistema de confiança* através de um *ecossistema de excelência*, a União Europeia, desejando proteger o *ser vivente*, elaborou – ainda em fase experimental –, um *know how to live with IA* em harmonia com as diretrizes do *supremo ser artificial social* (o Estado), e, mormente, os compromissos essenciais assumidos em múltiplas, porém convergentes, *Cartas de Direitos* para com seus tutelados, em um diálogo entre os arcabouços teóricos e práticos – até mesmo junto do setor privado e independente –, ou seja, um documento entrelaçador de políticas económico-sociais voltados ao desenvolvimento válido e seguro da *inteligência artificial*, não como um *novo lobo* do homem, mas cúmplice ao *futuro* cada vez mais *tecnológico* sob o *controlo* da pessoa.

4. Notas (*longe de ser*) conclusivas

Negar ao novo *ente artificial* a continuidade da sua presença, ainda, amistosa junto de nós, seres viventes em uma estrutura espaço-territorial sob a égide do Direito, como se conseguíssemos viver sem as utilidade dos mesmos ou, então, limitar-lhes a sua atuação apenas ao campo do lazer, do lúdico, jamais ao campo da ordem pública, seria eclipsar a rota natural do progresso da raça humana através das novas ferramentas que lhe promove segurança e independência. Entretanto, este novo tempo, a denominada *Era Digital*, incita a reflexão sobre as imputações que os *riscos* advindos do novo conteúdo aos textos originais dos direitos fundamentais, sendo imprescindível uma medida que possa conciliar o *futuro* e a *ordem pública*.

Nessa contínua releitura da vida, o *Parlamento Europeu e o Conselho* foram sábios ao apresentarem, no plano legislativo, uma Proposta de Regulamentação, cujo proceder traz um reconhecimento dos riscos e um meio de os isolar através de medidas que devem ser observadas para que sejam eles minimizados ou – quiçá – extirpados, e, consequentemente, seja erigido uma relação, antes de tudo, de confiança. Seguro é que a Diploma proposto representa o *início* da regulamentação, ensaiando equilibrar os dois lados da IA, isto é, o contributo promocional

---

– objeto de relações jurídicas. Em sentido similar, ao discutir sobre a responsabilidade da máquina, pronunciou-se Susana Aires de Sousa, «"Não fui eu, foi a máquina": teoria do crime, responsabilidade e inteligência artificial», 74-80.

de maior segurança à sociedade, ante a agilidade na resolução/prevenção de crimes, com os desafios inerentes aos sistemas que repercutem no âmbito da ordem pública – opacidade, viés discriminatório e violação da privacidade – através da enumeração e hierarquização dos *riscos, cujo movimento pendular vai do mínimo ao inaceitável*.

No *meio-termo* do *risco*, classificou-se aqueles que elevados eram e os requisitos que deveriam cumprir para o seu manuseio e funcionamento na máquina social, sem qualquer prejuízo às engrenagens humanas através das fricções e desgastes corrosivos pela privação da substância jurídico-fundamental – elemento garantidor de máxima exatidão do sistema amigo dos Direitos (dos) Humanos.

Portanto, por si só, a *Proposta do Regulamento* seja um pioneirismo, aguardando – decerto – novos acréscimos e/ou correções, especialmente por não haver, neste momento, conhecimento e pesquisa suficiente para afirmar se persistirá, embora elencados foram os requisitos protetores, a ameaça aos direitos fundamentais humanos, descortinando a opacidade sistémica e os vícios da discriminação, auxiliando a humanidade, com uma roupagem ética, em sua nova alvorada evolucional, não como *lobo*, mas o *cão* - animal a quem se considera o título de «melhor companheiro» – do homem.

Referências bibliográficas

Arendt, Hannah, *A condição humana*, 10ª ed., Rio de Janeiro: Forense Universitária, 2007.

Bachner, Jennifer, *Predictive policing: preventing crime with data and analytics*, Washington: IBM Center for the business of government, 2013, [Consult. 17 abril 2023]. Disponível em: http://www. businessofgovernment.org/sites/default/files/Predictive%20Policing.pdf.

Barbosa, Mafalda Miranda, *Inteligência Artificial: entre a utopia e a distopia, alguns problemas jurídicos*, Coimbra: Gestlegal, 2021.

Burchard, Christoph, «Artificial Intelligence as the End of Criminal Law? On the Algorithmic Transformation of Society» in Maria João Antunes / Susana Aires de Sousa, *Artificial Intelligence in the economic sector prevention and responsibility*, Coimbra: Instituto Jurídico Faculdade de Direito da Universidade de Coimbra, 2021, 165-205.

Canotilho, J. J. Gomes, *Direito constitucional e teoria da constituição*, 7ª ed., 11ª reimp., Coimbra: Almedina, 2003.

Criado, Natalia; Such, Jose M., «Digital discrimination», in Karen Yeung / Martin Lodge, *Algorithmic regulation*, 2019, 82-97.

Dolinger, Jacob, *A evolução da ordem pública no direito internacional privado*, Rio de Janeiro: Almedina, 1979.

Ehrhardt Jr., Marcos / Silva, Gabriela Buarque Pereira, «Diretrizes éticas para a inteligência artificial confiável na união europeia», *Jurismat* 12 (nov. 2020) 305-337;

Fetzer, James H, *Artificial Intelligence: Its scope and limits*, Kluwer Academic Publishers, 1990.

Floridi, Luciano, «Soft Ethics and the Governance od the Digitial», *Philosophy & Technology* 31/1 (mar. 2018) 1-8. [Consult. 17 abril 2023]. Disponível em: https://www.researchgate.net/publication/323248541_Soft_Ethics_and_the_Governance_of_the_Digital/link/5a895f23458515b8af92826f/download.

Floridi, Luciano, «Ethics after the information revolution», in Luciano Floridi, *The Cambridge handbook of information and computer ethics*, Cambridge: Cambridge University Press, 2010, 3-19.

Floridi, Luciano, *Philosophy and computing: an introduction*, London and New York: Routledge, 2001.

Hardyns, Wim/Rummens, Anneleen, «Predictive Policing as a New Tool for Law Enforcement? Recent Developments and Challenges», *European Journal on Criminal Policy and Research* 24 (2018) 201-218.

Heraclitus, of Ephesus, *Fragments: the collected wisdow of Heraclitus*, trad. Brooks Haxton, New York: Viking Penguin, 2011.

Kaplan, Jerry, *Artificial Intelligence: what everyone needs to know*, Oxford: Oxford University Press, 2016.

Merabet, Samir, *Vers un droit de l'intelligence artificielle*, Paris: Dalloz, 2020.

Pereira, Rui Soares, «Sobre o uso de sistemas de identificação biométrica (e de tecnologias de reconhecimento facial) para fins de segurança pública e de aplicação coercitiva da lei: reflexões a propósito da proposta de regulamento europeu sobre a inteligência artificial», *Revista da Faculdade de Direito da universidade de Lisboa* 63/1-2 (2022) 839-865.

Poças, Luís. *Manual de Investigação em Direito: metodologia da preparação de teses e artigos jurídicos*. 2ª ed, Coimbra: Almedina, 2022.

Prieto Álvarez, Tomás, *La dignidade de la persona: núcleo de la moralidad y el orden públicos, límite al ejercicio de libertades públicas*, Navarra: Editorial Aranzadi, 2005.

Ratcliffe, Jerry, «What is the future… of predictive policing?», *Translational Criminology* (spring 2014) 4–5;

Roa, Julio O. de, *Del orden publico em derecho positivo*, Buenos Aires: Librero Editor, 1926.

Rocha, Manuel Lopes, «Nota prévia», in Manuel Lopes Rocha / Rui Soares Pereira, coord., *Inteligência artificial & direito*, reimp., Coimbra: Almedina, 2022, 5-9.

Santos, António Pedro Ribeiro dos, *O estado e a ordem pública: as instituições militares portuguesas*, Lisboa: Instituto Superior de Ciências Sociais e Políticas, 1999.

Sieyès, Emmanuel-Joseph, *¿Qué es el estado llano?: ensayo sobre los privilegios*; versão de José Rico Godoy, Madrid: Centro de Estudios Constitucionales, 1988.

Sousa, Susana Aires de, «A IA no setor económico: uma reflexão entre o bom, o mau e o vilão», in Anabela Miranda Rodrigues, coord., *A inteligência artificial no direito penal II*, Coimbra: Almedina, 2022, 175-205.

Sousa, Susana Aires de, «"Não fui eu, foi a máquina": teoria do crime, responsabilidade e inteligência artificial», in Anabela Miranda Rodrigues, coord., *A inteligência artificial no direito penal*, Coimbra: Almedina, 2020, 59-93.

Sousa, Susana Aires de, *Neurociências e direito penal: em busca da verdade perdida (na mente)*, Coimbra: Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2017, [Consult. 17 abril 2023], Disponível em: https://www.uc.pt/site/assets/files/435430/direitonumahora2_neurocie_ncias_e_direito_penal.pdf.

Turing, Alan, *The essential Turing: seminal writings in computing, logic, philosophy, artificial intelligence, and artificial life plus the secrets of enigma*, editado por B. Jack Copeland, New York: Oxford University Press, 2004.

Zippelius, Reinhold, *Teoria geral do estado*, trad. de António Cabral de Moncada, 2ª ed., Lisboa: Fundação Calouste Gulbenkian, 1995.

Outros documentos

Comissão Europeia, *Livro branco sobre a Inteligência Artificial - uma abordagem europeia virada para a excelência e a confiança*, 2020, [Consult. 17 abril 2023], disponível em: https://op.europa.eu/pt/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1.

Parlamento Europeu/Conselho. *Resolução do Parlamento Europeu sobre a inteligência artificial no direito penal e a sua utilização pelas autoridades policiais e judiciárias em casos penais*, 2021, [Consult. 17 abril 2023], disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2021-10-06_PT.html.

Parlamento Europeu/Conselho. *Proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021, [Consult. 17 abril 2023], disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX:52021PC0206.

Parlamento Europeu/Conselho. *Anexos da Proposta de Regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da União*, 2021, [Consult. 17 abril 2023], disponível em: https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0004.02/DOC_2&format=PDF, 5-6

União Europeia. *Carta dos Direitos Fundamentais da União Europeia*, 2016, [Consult. 17 abril 2023], Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:12016P/TXT&from=FR.

União Europeia. *Regulamento geral sobre a proteção de dados da União Europeia* n. 679/2016, [Consult. 17 abril 2023], disponível em https://www.pgdlisboa.pt/leis/lei_mostra_articulado.php?nid=2961&tabela=leis.

ISBN: 978-989-9075-50-4

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR