UNIVERSIDADE D
COIMBRA

Olga Stelmashchuk

Algoritmos Preditivos na Justiça Penal

Uma análise à luz do Novo Regulamento da União Europeia sobre a Inteligência Artificial

Dissertação no âmbito do Mestrado em Ciências Jurídico-Forenses orientada pela Professora Doutora Susana Maria Aires de Sousa e apresentada à Faculdade de Direito da Universidade de Coimbra

Julho de 2023

FACULDADE DE DIREITO
UNIVERSIDADE D
COIMBRA

OLGA STELMASHCHUK

ALGORITMOS PREDITIVOS NA JUSTIÇA PENAL

Uma Análise à luz do Novo Regulamento da União Europeia sobre a Inteligência Artificial

PREDICTIVE ALGORITHMS IN CRIMINAL JUSTICE

An Analysis in light of the New European Union Regulation on Artificial Intelligence

Dissertação apresentada à Faculdade de Direito da Universidade de Coimbra no âmbito do 2º Ciclo de Estudos em Direito (conducente ao grau de Mestre), na Área de Especialização em Ciências Jurídico-Forenses

Orientadora: Professora Doutora Susana Maria Aires de Sousa

Coimbra, 2023

“Our intelligence is what makes us smarter,
and AI is an extension of that quality”.
– Yann Lecun

“Then I thought: what if I could play together with a computer at my side, combining our strengths? Human intuition plus machine's calculation, human strategy, machine tactics, human experience, machine's memory. Could it be the perfect game ever played?”
– Garry Kasparov

AGRADECIMENTOS

Findo este projeto que marca o término de uma etapa importante no meu percurso académico, cabe dirigir os sentidos agradecimentos às pessoas que até aqui me acompanharam.

O meu maior obrigada não podia deixar de ser para os meus pais, os principais impulsionadores desta caminhada e, de longe, aqueles que mais acreditam em mim. Sem vocês nada seria possível, a vocês tudo devo. Também umas palavras de apreço às minhas avós pelo carinho incondicional, o vosso aconchego torna tudo mais leve.

À minha irmã, pelos momentos de descontração e ânimo desta jornada. E aos meus amigos, pela motivação e palavras de apoio que nunca escassearam.

Por último, manifesto o sentimento de muita admiração à Dra. Susana Aires de Sousa por ser um exemplo de professora com a qual aprendi muito. Agradeço-lhe os conselhos e total disponibilidade que sempre demostrou durante a elaboração deste trabalho.

SUMÁRIO

Esta dissertação procurou analisar a interseção possível, atual e eventual, entre o direito e as novas tecnologia, considerando o ascendente potencial dos sistemas de Inteligência Artificial. Estes sistemas comprovaram superar consideravelmente certas capacidades humanas, com vantagens na precisão, rapidez e eficiência, na execução de diagnósticos, prognósticos e tarefas analíticas. No entanto, a sua aplicação no âmbito da justiça penal revela também vários inconvenientes. Tornou-se, assim, necessário identificar estes prós e contras que podem estar associados à utilização de instrumentos que têm por base algoritmos de IA, em particular, nos vários momentos em que estes se materializam em resultados e decisões automatizadas com potencial impacto na vida dos seus destinatários.

Para tanto, cingimos esta investigação ao estudo do fenómeno da justiça preditiva, em particular, no que diz respeito à utilização de instrumentos de previsão do risco na atividade policial e na administração da justiça. A introdução de instrumentos tecnológicos cada vez mais "inteligentes" no processo penal vai, por um lado, desafiar as conceções tradicionais da teoria da infração e, por outro, colocar sobre as instituições estaduais (e não só) uma pressão acrescida para agirem no sentido da contenção dos seus riscos, respondendo a uma necessidade de equilibrar o avanço tecnológico com a proteção dos direitos fundamentais das pessoas, sem prescindir dos benefícios que dele possam resultar.

Uma dessas respostas, vertida na primeira Regulamentação de Inteligência Artificial (Lei da UE sobre IA), tem ocupado a Comissão Europeia e o Parlamento Europeu nos últimos anos e entra agora na sua fase decisiva, a curtos passos da aprovação. Foi, portanto, na proposta deste diploma que procurámos algumas luzes sobre como será o futuro desta admirável tecnologia no espaço europeu.

Palavras-chave: algoritmos; inteligência artificial; direito penal; justiça preditiva; regulamento; União Europeia

ABSTRACT

This dissertation sought to analyse the possible, current, and eventual intersection between law and modern technologies, considering the ascending potential of Artificial Intelligence systems. These systems have proven to considerably surpass specific human capacities, with advantages in precision, speed, and efficiency, in the execution of diagnostics, prognostics, and analytical tasks. However, its application in criminal justice also reveals several drawbacks. It has therefore become necessary to identify these pros and cons that may be associated with using instruments based on AI algorithms, in the various moments in which these materialize in automated results and decisions with a potential impact on the lives of its recipients.

To this end, we limited this investigation to the study of the phenomenon of predictive justice, concerning the use of risk forecasting instruments in police activity and the administration of justice. The introduction of increasingly "smart" technological tools in criminal proceedings will, on the one hand, challenge the traditional conceptions of the theory of the offense and, on the other hand, place increased pressure on state institutions (and beyond) to act in the sense of containment of its risks, responding to a need to balance technological progress with the protection of people's fundamental rights, without ignoring the benefits that may result from it.

One of these responses, translated into the first Regulation on Artificial Intelligence (AI Act), has occupied the European Commission and the European Parliament in recent years and is now entering its decisive phase, just a few steps away from approval. Therefore, in the proposal of this diploma, we looked for some light on what the future of this admirable technology in Europe will be like.

Key-words: algorithms; artificial intelligence; criminal law; predictive justice; regulation; European Union

LISTA DE SIGLAS E ABREVIATURAS

Ac.	– Acórdão
AI	– *Artificial Intelligence*
AGI	– *Artificial General Intelligence*
al.	– alínea
art./arts.	– artigo/s
CAS	– *Crime Anticipation System*
CEPEJ	– Comissão Europeia para a Eficiência da Justiça
Cf.	– Confira/conforme
COMPAS	– *Correctional Offender Management Profiling for Alternative Sanctions*
Consid.	– Considerando
Coord/coords.	– Coordenador/es
CP	– Código Penal
CPC	– Código de Processo Civil
CPP	– Código de Processo Penal
CRP	– Constituição da República Portuguesa
EBS	– *Evidence-based sentencing*
Ed.	– Edição
Eds.	– Editores
e.g.	– *exempli gratia* (“por exemplo”)
et. al.	– *et alia* (“e outros”)
EUA	– Estados Unidos da América
GPT	– *Generative Pre-Treined Transformer*
HART	– *“Harm Assessment Risk Tool”*
IA	– Inteligência Artificial
Ibid.	– *Ibidem* (“no mesmo lugar”)
ICSE	– *International Child Sexual Exploitation*
i.e.	– “isto é”
IoT	– *Internet of Things*
IRIS	– “Informação, Racionalização, Integração e Sumarização”
Loc. cit.	– no lugar citado (“na mesma página”)

ML	– *Machine Learning*
n.º	– número
OPC	– Órgãos de Polícia Criminal
Op. cit.	– obra citada
p./pp.	– página/páginas
RGPD	– Regime Geral de Proteção de Dados
Ss.	– seguintes
STJ	– Supremo Tribunal de Justiça
SyRi	– *System Risk Indication*
Trad.	– tradução
UE	– União Europeia
v.	– versus
Vol.	– volume
XAI	– *"Explainable Artificial Intelligence"*

ÍNDICE

I. INTRODUÇÃO

A Quarta Revolução Industrial, nome utilizado para descrever a fase atual de transformação da indústria (também conhecida como Indústria 4.0), impulsionada pela adoção de recursos digitais e de conectividade como a robótica, a Inteligência Artificial (doravante IA) e a Internet das Coisas (*Internet of things – IoT),* baseia-se na infraestrutura que adveio da revolução digital e veio tornar as tecnologias mais sofisticadas e integradas[1]. Trata-se de uma tendência global anunciada como uma transformação sem precedentes, tanto na sua amplitude e profundidade, como na velocidade da inovação[2]. Este paradigma tem-se vindo a massificar com o acesso ao crescente volume de dados disponíveis (*Big Data*) e o aumento da capacidade de processamento dos novos computadores que podem ser treinados para cumprir tarefas específicas e encontrar padrões numa imensidão de dados.

Num advento de uma sociedade marcadamente tecnológica, estas inovações já revolucionaram a maneira como vivemos, nos relacionamos e trabalhamos – implementadas em carros autónomos, assistentes virtuais, softwares de tradução, reconhecimento facial –, mas também se vão progressivamente entranhando nas áreas mais críticas da sociedade como os setores da indústria, da saúde, (ciber)segurança e economia. Não sendo o direito alheio a este fenómeno, é um dos campos onde os avanços da IA, para além de grandes oportunidades, podem trazer também diversos riscos, principalmente no direito penal.

A transição para uma sociedade digital resultou em transformações significativas, "não só quanto ao contexto em que os crimes ocorrem, mas também quanto ao modo como a investigação é feita", bem como a forma como as prova são produzidas e as decisões ao longo do processo são tomadas[3]. As novas tecnologias, as aplicações de IA e a robótica em particular, surgem, assim, como uma "faca de dois gumes" que, por um lado, permite melhorar o *modus operandi* dos grupos criminosos, ou criar outros tipos legais de crime[4] –

[1] Cf. SCHWAB, Klaus, (2016). *A Quarta Revolução Industrial*. Trad. Daniel M. Miranda. São Paulo: Edipro.
[2] *Ibid.*
[3] RODRIGUES, Anabela M., (2020a). "A justiça preditiva entre a americanização e a europeização", p. 12, in RODRIGUES, Anabela M. (Coord.), (2020). «A Inteligência Artificial no Direito Penal», Coimbra: Almedina.
[4] Cf. INTERPOL/UNICRI, (2019). *"Artificial Intelligence And Robotics For Law Enforcement"*, pp. 5 e ss., que, neste sentido, refere os crimes *high-tech* e à necessidade de o direito estar preparado para alavancar novas tecnologias para melhor prevenir e controlar o crime. O sentido dual da inovação também se reflete na possibilidade de a mesma tecnologia ser empregue para, por um lado, aumentar a eficiência de certas práticas, e, por outro, para cometer crimes, veja-se o exemplo da negociação algorítmica de alta frequência (*high frequency trading*) usada na compra e venda de ações, mas também aplicada em estratégias de manipulação da informação de mercado como *spoofing*. Cf. SOUSA, Susana A. de, (2020a). ""Não fui eu, foi a máquina":

num fenómeno de "alargamento das manchas de criminalização"[5] (*e.g.*, o cibercrime[6]) –, e, por outro, pode introduzir "grandes oportunidades no domínio da execução da lei, em particular na melhoria dos métodos de trabalho das autoridades policiais e judiciais"[7] no sentido de tornar mais eficiente a prevenção e combate a essa mesma criminalidade. Por isso, os agentes responsáveis pela aplicação da lei já não operam apenas no paradigma de punição *ex post facto*, mas procuram, cada vez mais, servir-se de medidas preventivas *ex ante*[8] no combate ao crime, o torna o recurso a algoritmos preditivos e ferramentas de IA sucessivamente mais apelativo. Esta configuração do crime, como um risco a ser calculado, leva uma sociedade pós-crime, centrada tradicionalmente na teoria do dano, onde as práticas de ordenamento surgem *post hoc*, a ceder lugar a uma sociedade pré-crime focada na teoria de risco, onde as práticas de ordenação são essencialmente preventivas[9]. Assim, uma atuação cada vez mais preocupada com a prevenção do crime que ainda não ocorreu (e pode nunca ocorrer) afeta, em última instância, a própria "essência da justiça penal, com a sua possível transformação numa justiça preditiva"[10] onde o direito penal se passa a centrar na probabilidade e na perigosidade do agente, determinadas com recurso a IA, e não no facto praticado.

Também a utilização de algoritmos inteligentes nos processos de tomada de decisão judicial promete oferecer uma proteção mais eficaz e eficiente dos interesses legais, bem como uma aplicação da lei mais neutra, objetiva e coerente[11] em comparação com os decisores humanos. Porém, a ideia que de antemão se perspetiva é que com grandes promessas de eficiência vêm associados grandes riscos. Ao longo deste trabalho procuraremos refletir sobretudo quanto ao seu impacto no domínio penal, os prováveis

---

teoria do crime, responsabilidade e Inteligência Artificial", p. 69, in RODRIGUES, Anabela M. (Coord.), (2020). «A Inteligência Artificial no Direito Penal», Coimbra: Almedina.

[5] RODRIGUES, Anabela M., (2020c). "A Política Criminal no Estado de Direito do Século XXI – os desafios da segurança". *Revista Brasileira de Ciências Policiais*, Vol. 11, n.º 1, p. 29.

[6] Regulado entre nós na Lei n.º 109/2009, "Lei do Cibercrime".

[7] (2020/2016(INI)) – "A Inteligência Artificial no direito penal e a sua utilização pelas autoridades policiais e judiciárias em casos penais", p. 4.

[8] ZAVRŠNIK, Aleš, (2021). "Algorithmic justice: Algorithms and big data in criminal justice settings". *European Journal of Criminology*, Vol. 18(5), p. 627.

[9] ZEDNER, Lucia, (2007). "Pre-crime and post-criminology?" *Theoretical criminology*, 11.2, p. 262.

[10] SOUSA, Susana A. de, (2020b). "Um Direito Penal desafiado pelo desenvolvimento tecnológico: alguns exemplos a partir das neurociências e da inteligência artificial", *Revista da Defensoria Pública da União*, n.º 14, p. 33.

[11] BURCHARD, Christoph, (2021). "Artificial Intelligence as the End of Criminal Law? On the Algorithmic Transformation of Society", pp. 181 e ss., in ANTUNES, Maria J. & SOUSA, Susana A. de, (Eds.), «Artificial Intelligence in The Economic Sector Prevention and Responsibility», *Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra.*

efeitos nefastos no plano dos direitos fundamentais dos visados e os possíveis conflitos com alguns princípios processuais.

Para isso, começaremos por percorrer algumas reflexões sobre o estado da arte da IA na conjuntura atual, em forma de uma pré-compreensão da tecnologia existente, passando pelos seus contornos, conceito(s) e as várias perspetivas que se abrem na sua utilização, em especial no âmbito da justiça criminal, por referência a dois "momentos" distintos da sua possível aplicação: (1) antes do processo, com o foco nos instrumentos de policiamento preditivo, e (2) durante o processo, onde a atenção incidirá nas ferramentas preditivas destinadas a fazer juízos de prognose e a facilitar a tomada de decisões. Nesta última hipótese, a tónica recairá sobre as decisões automatizadas, i.e., decisões judiciais tomadas por intermédio de softwares, e a sua admissibilidade durante as várias fases do processo. Aqui procuraremos chegar a um ponto de reflexão sobre a possibilidade de, um dia, admitirmos juízos produzidos por máquinas "inteligentes" no processo decisório dos julgadores penais, correndo, por fim, o risco de deixarmos de ter uma decisão judicial em detrimento de uma decisão científica ou técnica, concretizada num panorama penal onde sistemas programados para atuarem de forma mais ou menos autónoma assumem o encargo da decisão.

Impõe-se aqui interrogar sobre a natureza que reveste os dados emitidos pelo algoritmo. Poderão as ferramentas de IA intervir na apuração dos factos em tribunal? Valerão os dados algorítmicos como prova? Poderá o decisor discordar do resultado sugerido? Quais as garantias ao dispor do visado contra a "análise" da máquina? Se é certo que se abrem mais questões do que as respostas que aqui poderemos explorar, partiremos da análise da experiência consolidada de ordenamentos jurídicos estrangeiros onde os sistemas preditivos de IA já são uns aliados da justiça, seja no combate à criminalidade, ou na atividade jurisdicional, e, a partir daí, analisaremos as principais objeções apontadas à aplicação de algoritmos de avaliação do risco em contexto criminal, sobretudo na fase judicial. Por último, não esqueceremos de enquadrar os esforços perpetuados pela União Europeia a caminho da regulação da IA e a sua posição mais recente à data deste trabalho de forma a limitar fronteiras até onde estes sistemas podem ser inofensivos (e até benéficos para a realização da justiça) e onde começa a lesão não permitida de direitos fundamentais, transgressora de princípios estruturantes do estado de direito.

II. O JOGO DA IMITAÇÃO

As soluções que até há pouco eram consideradas conceções futurísticas, alimentadas por cenários distópicos da literatura e do cinema, saíram do plano da fantasia e são cada vez mais a realidade do "hoje". Porém, ao contrário do ficcionado em filmes disruptivos como "Eu, Robô" e "*Ex Machina*", onde a IA é personificada por robôs humanoides que dominam o mundo, a ascensão das tecnologias atuais de IA, por mais impressionante que seja, está longe de ser assim tão sofisticada.

Embora a origem da tecnologia necessária para o seu desenvolvimento remontar aos tempos da Segunda Guerra Mundial, o termo "Inteligência Artificial" que conhecemos hoje apenas surgiu anos depois, através das pesquisas do matemático John McCarthy[12]. Antes disso, o cientista Alan Mathison Turing, conhecido como o "pai da computação", já estudava o conceito de "inteligência mecânica" desde, pelo menos, 1941, e, em 1947, falava publicamente em "inteligência computacional".[13]

Em 1950, Turing já tinha apresentado à comunidade académica um estudo em que concluía que as máquinas poderiam ser programadas para aprender pelo mimetismo da inteligência humana[14]. Para provar isso, criou um teste baseado no tradicional "Jogo da Imitação"[15]. Neste exercício, uma pessoa, um computador e um observador humano (juiz) são colocadas em salas diferentes e só podem comunicar por texto dactilografado. Um computador passa no teste e é considerado inteligente, se o observador, depois de analisar o conteúdo da conversa desenvolvida entre a pessoa e a máquina, não consegue distinguir quem é quem pelas respostas apresentadas. Desde então, o teste inspirou várias discussões

[12] Apresentadas durante o seminário "*Dartmouth Summer Research Project on Artificial Intelligence*" (1956) que se tornou conhecido como o evento fundador da IA como uma área de estudo autónoma. Neste projeto os cientistas procuraram estudar como poderiam fazer as máquinas resolverem tipos de problemas até então reservados aos humanos, tendo por base a hipótese de que a aprendizagem ou qualquer outra característica humana podia, em princípio, ser descrita com tanta precisão que seria possível uma máquina simulá-la. Cf. MCCARTHY, John, *et al.*, (1955). A proposal for the Dartmouth summer research project on artificial intelligence. http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

[13] COPELAND, B. Jack, (2000). The Turing Test. *Minds and Machines,* 10, pp. 519–539. https://doi.org/10.1023/A:1011285919106

[14] TURING, Alan M., (1950). "Computing Machinery and Intelligence", Mind, *A Quarterly Review of Psychology and Philosophy*, Vol. LIX, n.º 236, pp. 433–460. https://doi.org/10.1093/mind/LIX.236.433.

[15] Deste jogo surge a intenção de identificar o capítulo explicativo sobre a IA com o mesmo nome, dada a sua propensão para replicar o comportamento humano. Um filme com o mesmo título foi vencedor do prémio Oscar,– *"The Imitation Game"* – baseado na vida de A. M. Turing e o seu papel decisivo no avanço das ciências da computação durante a Segunda Guerra Mundial ao desenvolver o aparelho eletromecânico "Bombe", um sistema de descodificação avançado capaz de decifrar a criptografia nazi gerada pela máquina "Enigma".

sobre a natureza da consciência e continua a ser um tópico debatido na filosofia e outras áreas que exploram a relação entre a mente e a máquina[16].

O filósofo John Searle, no seu artigo "Minds, Brains, and Programs"[17], criticou a metodologia o Teste de Turing alegando que este não podia provar realmente a inteligência de uma máquina porque os sistemas, ainda que passem no teste, nunca terão a consciência para compreender o que está a ser discutido.

Para sustentar a sua objeção, Searle apresentou o argumento do "Quarto Chinês" como uma analogia para as limitações de uma IA que se baseia no processamento de informações simbólicas. A experiência consiste em colocar uma pessoa que não fala chinês num quarto com um conjunto de instruções para manipular os símbolos chineses. Essas regras vão permitir que a pessoa responda por escrito às perguntas em chinês, embora ela não entenda o significado dos caracteres[18]. Ao responder corretamente às perguntas colocadas em chinês, ela passa no teste de Turing, porém, não significa que ela saiba chinês. Por conseguinte, de acordo com Searle, a execução do programa correto (isto é, apresentação dos *outputs* certos) não gera necessariamente a "compreensão" nem um comportamento consciente[19]. Esta crítica originou também a divisão entre os conceitos de IA "Fraca" e IA "Forte".

A Inteligência Artificial Fraca ou Restrita (Weak-AI/Narrow-AI) consiste num sistema computacional que é projetado para resolver ou completar tarefas específicas[20] com precisão dentro de um contexto pré-definido pelo ser humano. Trata-se de uma forma limitada de IA sem capacidade cognitiva independente e sem possibilidade de operar fora do contexto específico para o qual foi programada. Na prática, pode auxiliar no processamento de um grande volume de informações, servindo para a classificação de dados, o reconhecimento da fala e de imagens, jogar xadrez, simulação de uma conversa nos *chatbots* ou a condução de um carro. Dada o amplo espectro de possibilidades, é neste campo que atualmente se verificam os maiores avanços.

---

[16] Sobre o teste e as objeções tecidas à proposta de Turing cf. OLIVEIRA, Arlindo, (2019). *Inteligência Artificial*, Fundação Francisco Manuel dos Santos, pp. 42-50.
[17] SEARLE, John R. (1980). "Minds, Brains and Programs", In: The Behavioral and Brain Sciences, 3, Cambridge: Cambridge University Press, pp. 417-457.
[18] OLIVEIRA, (2019), p. 85.
[19] *Ibid;* SEARLE, (1980).
[20] HAENLEIN, Michael; KAPLAN Andreas, (2019). "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence", Business Horizons, Vol. 62(1), p. 16. https://doi.org/10.1016/j.bushor.2018.08.004

Por sua vez, na Inteligência Artificial Forte (Strong-AI)[21], "os computadores com os programas certos poderiam estar, literalmente, preparados para compreender e ter outros estados cognitivos"[22]. Esta possibilidade, sendo meramente hipotética para a maioria dos investigadores, existe apenas num plano teórico e descreve um certo nível de desenvolvimento de IA capaz de se aproximar da mentalidade humana numa ampla variedade de tarefas cognitivas, incluindo atividades intelectuais complexas, como o raciocínio, o planeamento, a autoconsciência[23]. Neste contexto, a referência será sempre a programas ou instrumentos de IA fraca que dependem da interferência humana tanto para definir os parâmetros dos seus algoritmos de aprendizagem como para fornecer os dados de treino relevantes para garantir a sua precisão, pois é, justamente, aquela que se utiliza "para estabelecer predições de decisões nos processos"[24].

Inicialmente, a definição de Inteligência Artificial era bastante simples e limitada. No entendimento de John McCarthy, seria "a ciência e a engenharia de criar máquinas inteligentes, especialmente programas de computador inteligentes"[25]. Hoje sabemos que, por mais sofisticados que sejam, nem todos os programas computacionais dão origem à IA. A maioria dos programas que podem ser considerados "inteligentes" hoje, i.e., dotados de "*machine intelligence*", dependem de tecnologias como o *machine learning*[26], carecem de

---

[21] A par da IA forte, é comum encontrar-se o termo "Inteligência Artificial Geral" ou *"Artificial General Intelligence"(AGI).* No entanto, estes nem sempre são usados como sinónimos, pois alguns autores utilizam o termo "IA forte" numa aceção muito próxima da noção que atualmente define a IA de terceira geração, chamada de "Superinteligência Artificial", que seria reservada para os sistemas verdadeiramente autoconscientes que, de certa forma, tornariam os humanos redundantes. Já o termo de AGI seria colocado num nível intermédio, uma IA de segunda geração, utilizado num sentido mais prático e técnico para definir sistemas capazes de raciocinar, "resolver problemas de forma autónoma e superar ou igualar-se aos humanos em várias áreas". Cf. HAENLEIN & KAPLAN, (2019), p. 16.

[22] SEARLE, (1980), p. 417. Tradução livre.

[23] SILVA, Susana C. e, (2021). "A Inteligência Artificial fraca e a fraca Inteligência Artificial", *Jornal Económico*. Disponível em: https://jornaleconomico.pt/noticias/a-inteligencia-artificial-fraca-e-a-fraca-inteligencia-artificial-764118/

[24] GUIMARÃES, Rodrigo R. C., (2019). "A Inteligência Artificial e a disputa por diferentes caminhos em sua utilização preditiva no processo penal". *Revista Brasileira De Direito Processual Penal*, Porto Alegre, Vol. 5, n.º 3, set./dez, pp. 1563-1564. Https://Doi.Org/10.22197/Rbdpp.V5i3.260

[25] Desta forma, definia-se o objetivo da AI com base no conceito de inteligência que seria "a parte computacional da capacidade de alcançar objetivos no mundo. Vários tipos e graus de inteligência ocorrem em pessoas, muitos animais e algumas máquinas." MCCARTHY, John, (2007). «*What is AI?/Basic Questions*», University of Standford, p. 2-3. Trad. livre. http://jmc.stanford.edu/articles/whatisai.html

[26] Sobre o funcionamento e técnicas de *machine learning (ML)* ver https://www.ibm.com/topics/machine-learning. Simplificadamente, ML é um subcampo da IA que "ensina" à máquina como ela deve aprender. Os modelos de ML atuam dentro de um campo de dados disponíveis, procuram por padrões e tentam tirar conclusões sem serem especificamente programados. "O criador do algoritmo descreve-lhe o problema a resolver e proporciona-lhe o acesso à informação, sendo que, a partir daí, o algoritmo, autonomamente, "aprende" a resolver a questão que inicialmente lhe foi colocada e resolve-a, sem intervenção humana." Cf.

acesso a uma enorme quantidade de dados[27] *(Big Data)* e são "capazes de analisar o seu ambiente e tomar medidas com um determinado nível de autonomia para atingir objetivos específicos"[28].

"O termo que se tornou parte da linguagem quotidiana abrange uma ampla variedade de ciências, teorias e técnicas cujo objetivo é fazer com que uma máquina reproduza as capacidades cognitivas de um ser humano"[29] como a "aprendizagem, adaptação, interação, raciocínio, resolução de problemas, representação de conhecimento, previsão e planeamento, autonomia, perceção, movimento e manipulação"[30].

Sendo, portanto, um conceito científico vasto, dificilmente conseguiríamos reduzi-lo a uma definição unânime, tanto que, até o momento, não existe uma formulação única de IA aceite pela comunidade científica que seja capaz de abranger toda a complexidade e técnicas utilizadas. Além disso, qualquer tentativa de definição deve procurar debruçar-se sobre as principais características distintivas da IA, de modo a diferenciá-la de outros sistemas de software ou de programação computacional mais simples.

Contudo, não se buscando fazer uma análise exaustiva deste conceito, destacamos apenas que atualmente são várias as definições de IA encontradas na literatura científica e que, segundo Russell e Norvig, podem ser agrupadas em 4 categorias: 1 – sistema que pensa como os seres humanos; 2 – sistema que age como seres humanos; 3 – sistema que pensa racionalmente; e 4 – sistema que age racionalmente[31]. Algumas das definições existentes medem o sucesso da IA com base em sua capacidade de replicar o desempenho humano,

LOBO, Fernando R., (2021). "A utilização de sistemas preditivos de inteligência artificial na justiça", Lusíada, Revista de Direito, 23/24, p. 50.

[27] *Ibid, Loc. Cit.*

[28] A Comissão Europeia enumera alguns exemplos destes sistemas: os que estão confinados ao software, atuando no mundo virtual – assistentes de voz, programas de análise de imagens, motores de busca, sistemas de reconhecimento facial e de discurso – e os que podem ser integrados em dispositivos físicos – robôs avançados, automóveis autónomos, veículos aéreos não tripulados ou aplicações da Internet das Coisas. Cf. COM (2018) 237 final. Bruxelas, 25.04.2018.

[29] CAHAI (2020) 23, Council of Europe. Feasibility Study, Ad hoc Committee on Artificial Intelligence, p. 3, Trad. livre.

[30] AMA, (2020). *Guia para a Inteligência Artificial*, p. 10. Disponível em https://tic.gov.pt/areas-tematicas/inteligencia-artificial

[31] RUSSELL, Stuart & NORVING, Peter, (2013). *Inteligência Artificial*, 3ª ed., trad. Regina C. S. de Macedo, Rio de Janeiro: Campus Elsevier.

enquanto para outras a medida de desempenho ideal é a chamada racionalidade – "um sistema é racional se faz a "coisa certa", dado o que ele sabe"[32].

No caso de "pensar" ou "agir como um humano" a abordagem é dominada pelo desejo de comparar ou aproximar a máquina a uma pessoa, tal como acontecia na proposta do teste de Turing. A base dessa abordagem é a compreensão de como uma pessoa atua ou pensa. Embora ainda esteja longe de alcançar a inteligência humana, como supra notado, muitas pesquisas em IA são inspiradas no estudo da mente humana e do pensamento racional. No entanto, até agora, a ciência moderna não foi capaz de dar uma explicação inequívoca sobre como se processam os mecanismos do pensamento na mente humana e muito menos como parametrizá-los. A tecnologia que mais se aproxima do funcionamento do cérebro, e da forma como ele assimila a informação, baseia-se em redes neuronais artificiais que são usadas nos mecanismos de aprendizagem profunda (*deep learning* enquanto técnica de *machine learning*) e inspiram-se no processo de sinapse que ocorre nos neurónios, ainda que, no estado atual da tecnologia, se encontrem longe de o recriar[33].

A própria questão de tentar compreender a natureza humana remonta a Aristóteles que a descrevia como algo essencialmente ligado à razão e à capacidade de pensar. E a IA, ao ambicionar replicar certos aspetos do comportamento e cognição humana, transpõe para o campo da IA certos princípios do raciocínio aristotélico, como os silogismos, para processar a informação e inferir conclusões lógicas[34].

Silogismo, do grego "*argumento*", representa o "raciocínio feito a partir de duas proposições (premissas), das quais se deduz uma terceira, a conclusão"[35]. O algoritmo, definido como uma "sequência lógica e finita de instruções que visam atingir um determinado propósito"[36], também funciona dessa forma. Mas apenas o silogismo científico[37], baseado na dedução e argumentação lógica, está ao alcance da análise realizada

[32] *Ibid*.
[33] OLIVEIRA, (2019), pp. 65-69.
[34] Sobre os fundamentos da inteligência artificial *vide* RUSSELL & NORVING, (2013).
[35] Porto Editora – *silogismo* no Dicionário infopédia da língua portuguesa. Porto: Porto Editora. Disponível em https://www.infopedia.pt/dicionarios/lingua-portuguesa/silogismo
[36] AMA, (2020). "Guia para a Inteligência Artificial", p. 9. Disponível em https://tic.gov.pt/areas-tematicas/inteligencia-artificial
[37] Por sua vez, a capacidade de demonstrar senso crítico, tecer uma opinião ou argumento fundamentado, criar um discurso estruturado com o objetivo de persuadir ou convencer o outro (silogismo dialético e retórico), continua reservada ao Homem e muito dificilmente será replicável num código algorítmico.

por softwares de IA "fraca", uma vez que só a lógica por trás deste silogismo oferece um "alto grau de previsibilidade" e permite saber qual será o resultado alcançado com base numa "determinada configuração algorítmica".[38]

Todavia, apesar de todas as aproximações e comparações com as características humanas, principalmente com a inteligência que, por si só, é de complexa explanação, é inegável que a IA tem necessariamente uma natureza diversa e, por isso, deve ser considerada como tal, "uma inteligência expressa em formas diferentes daquelas que estão normalmente associadas a um ser humano"[39]. Ora, não sendo esta "inteligência" (ainda) capaz de substituir totalmente o pensamento humano, analisaremos as possibilidades de o complementar ou auxiliar, no que para aqui interessa, nos momentos de ponderação decisória nas atividades forenses.

[38] GUIMARÃES, (2019), p. 1569.
[39] SILVA, (2021).

III. INTELIGÊNCIA ARTIFICIAL NO DIREITO

A inovação digital e a introdução de IA, em particular, estão na agenda do dia dos vários países desenvolvidos. A competitividade e a necessidade de dar respostas cada vez mais céleres e eficazes, tem acelerado a corrida às ferramentas de IA não só dos setores privados como também públicos. A sua inclusão em áreas como a saúde, finanças, educação, segurança, economia e justiça tem sido uma prioridade, indo de encontro aos Objetivos de Desenvolvimento Sustentável das Nações Unidas. Atualmente, amplamente difundida, a IA já é aplicada em diversos planos. Os exemplos[40] vão desde os veículos autônomos (como drones e automóveis autónomos), diagnóstico médico, sugestões de entretenimento e criação de conteúdo "à medida" (copywriting, poesia, música e mesmo filmes[41] ou imagens[42]), prova de teoremas matemáticos, jogos (como o xadrez ou *Go*), filtragem de spam, marketing online, assistentes de voz (como a *Siri* ou *Alexa*), até ao reconhecimento e previsão comportamental (previsões de tempo, de padrões de consumo, alterações no mercado de ações e previsão de decisões judiciais).

1. A DIGITALIZAÇÃO DA JUSTIÇA

No que ao direito diz respeito, assistimos a países empenhados na construção de uma Justiça Eletrónica onde a transformação digital assume-se como um eixo estratégico a prosseguir. Esta Justiça Eletrónica, ou *e*-Justiça, caracteriza-se pela adoção de "tecnologias da informação e comunicação em tarefas do âmbito da justiça, tais como simplificação e desmaterialização de processos judiciais (…) e eliminação e simplificação de atos e procedimentos" que tramitam nos tribunais[43].

---

[40] Enumerados em NOVAIS, Paulo; FREITAS, Pedro M., (2018). «Inteligência Artificial e Regulação de Algoritmos», *Diálogos, União Europeia-Brasil*, maio, pp. 15-16.

[41] Em novembro de 2022, a empresa OpenAI lançou o "ChatGPT" (sigla para "Generative Pre-Trained Transformer") – uma aplicação movida a IA generativa capaz de criar conteúdo escrito sobre qualquer assunto. Trata-se de uma plataforma de conversação que se vai aperfeiçoando através dos diálogos desenvolvidos. Uma vez que responde com base na informação que encontra na Internet, o programa tem as suas limitações, pelo que pode gerar respostas falsas ou imprecisas, bem como levantar preocupações sobre os direitos de autor e plágio. Em março de 2023, surgiu o GPT-4, uma versão avançada com a novidade de processamento de imagem e criação de vídeos que veio ampliar o âmbito de impacto desta tecnologia. Para mais detalhe: https://www.sciencefocus.com/future-technology/gpt-3/

[42] O sistema "DALL-E", também desenvolvido pela OpenAI, chama a atenção pela sua capacidade de gerar ilustrações, designs e gráficos a partir de qualquer texto. https://shifter.pt/2021/01/dall-e-gpt3-openai/

[43] https://apdsi.pt/glossario/j/justica-eletronica/

O fenómeno da digitalização na justiça portuguesa, confirma essa opção estratégica. No relatório "Transformação Digital Da Justiça 2015-2022" apresenta-se o balanço das políticas na área governativa da Justiça e destacam-se os projetos implementados para promover a centralização no utilizador dos serviços de justiça, numa lógica de *"digital by default",* inclusive através do aproveitamento das possibilidades oferecidas pelas novas tecnologias e soluções de IA. Atualmente em curso, a terceira edição do plano "*Justiça+Próxima*" dá continuidade ao processo de modernização, prolongando a sua implementação até 2025, com várias medidas que pretendem tornar o sistema judiciário mais transparente, acessível e eficaz – nomeadamente através de "esforços no desenvolvimento de programas com recurso à inteligência artificial para apoio às decisões dos tribunais (…) tendo também presente que o juiz será sempre o decisor último de qualquer processo."[44]

A nível europeu, no que concerne à utilização de tecnologia pelos tribunais e serviços do Ministério Público, "Portugal é o sexto país da União Europeia com maior digitalização do sistema judiciário", segundo dados de 2022, atendendo a indicadores como sistemas de comunicação à distância, "gestão eletrónica de processos, distribuição automática, teletrabalho para juízes e outros oficiais de justiça e ainda recurso a aplicações de IA".[45]

2. AS APLICAÇÕES DE IA NO CAMPO JURÍDICO

O estudo dos artigos científicos, bem como da prática jurídica estrangeira, permite-nos concluir que não se trata de uma introdução totalmente nova. O quotidiano forense já tem beneficiado da utilização de ferramentas movidas a IA, muito graças às assim chamadas *legaltech start-ups* que, cientes do potencial que a tecnologia tem para impactar as nossas vidas, têm investido no desenvolvimento de plataformas tecnológicas que facilitam o acesso aos serviços jurídicos e o seu funcionamento.

---

[44] Transformação Digital da Justiça 2015-2022, jan. 2022. Disponível em https://justica.gov.pt/Transformacao-Digital-da-Justica-2015-2022. Também no âmbito da política nacional para as Competências Digitais, foi apresentada a Estratégia Nacional de Inteligência Artificial – *«AI Portugal 2030»,* com vários eixos de ação, que pretende promover a investigação e a inovação, em prol do seu desenvolvimento.
Cf. https://www.incode2030.gov.pt/aip-2030/

[45] Notícia em: https://www.tsf.pt/portugal/sociedade/portugal-e-o-5-pais-da-ue-com-maior-duracao-dos-processos-nos-tribunais-16496519.html

Isto tem sido especialmente relevante, no âmbito das profissões jurídicas, desde logo, em tarefas que permitem a automação. A distribuição e gestão eletrónica de processos[46] e documentos, bem como a possibilidade de realizar um julgamento remotamente, já foi introduzida no âmbito processual de muitos sistemas judiciais. Outras áreas de aplicação que podem beneficiar de sistemas de IA remetem a tarefas mais rotineiras e repetitivas como: a análise ou tradução de documentos, gestão de prazos e notificações[47], pesquisas de legislação e jurisprudência, aconselhamento jurídico personalizado através da apreciação dos elementos objetivos segundo a legislação aplicável; previsão das decisões judiciais, traçar o perfil decisório de juízes; estudo dos possíveis cenários e probabilidades de êxito de uma ação judicial, numa lógica de custo/benefício, através da análise de precedentes e identificação de padrões, entre outras.

Cada vez mais acessíveis e praticamente à distância de um *click*, são várias as plataformas que oferecem assistência legal com ajuda da tecnologia de IA, não apenas para juristas, mas para qualquer interessado que procure lidar com assuntos jurídicos do dia a dia. Muitos dos recursos consistem em avançados motores de pesquisa, enquanto outros chegam mesmo a sugerir soluções práticas. A aplicação "*DoNotPay*" impugna multas de estacionamento, redige reclamações e aconselha sobre os direitos dos consumidores[48]. "*Westlaw*" (Thomson Reuters Legal) é um serviço de pesquisa com várias ferramentas de redação e revisão de documentos jurídicos[49]. "*Lex Machina*" (LexisNexis) – plataforma de análise jurídica que permite aos escritórios de advocacia e empresas preverem os resultados das diferentes estratégias jurídicas[50]. Também a IBM[51] chegou a apresentar o sistema

[46] Entre nós regulada no art.º 204.º no Código de Processo Civil (CPC).
[47] Uma citação ou notificação dos atos processuais realizados de forma automatizada por meio de sistemas de IA, mediante o cruzamento prévio da morada constante dos autos com as moradas registadas nas bases de dados oficiais ou por transmissão automática para o endereço eletrónico do citando, sem necessidade de intervenção humana, "permitiria garantir um melhor aproveitamento dos recursos, bem como uma maior eficiência no que concerne à tramitação e à gestão do processo judicial." GONÇALVES, Marco C., (2022). "Inteligência artificial e processo judicial: em busca da celeridade, da eficiência e da qualidade da justiça", pp. 272-274, in SILVA, Eva S. M. da & FREITAS, Pedro M. (Coords.), (2022). «Inteligência Artificial e Robótica: Desafios para o Direito do Século XXI» 1.ª Ed., Gestlegal.
[48] https://donotpay.com/
[49] https://legal.thomsonreuters.com/en/westlaw
[50] https://lexmachina.com/about/
[51] A IBM Research, empresa estadunidense conceituada na área da tecnologia e informática, ficou conhecida por desenvolver o supercomputador IBM Deep Blue que venceu o melhor jogador de xadrez do mundo, Gary Kasparov, numa partida que ficou conhecida como "*the Man vs. the Machine*". Mais tarde, em 2011, alcançou um novo sucesso quando outro computador, o IBM Watson, baseado numa tecnologia de computação cognitiva, venceu os dois campeões no programa televisivo "*Jeopardy!*", um jogo de conhecimentos gerais.

"*Ross*", o robô advogado que permitia[52] não só acompanhar todas as atualizações de legislação e precedentes, como oferecer soluções analíticas para as questões legais que lhe eram colocadas[53]. Em Portugal, aplicações de IA ainda têm pouca expressão prática, podendo, ainda assim, serem encontradas em alguns portais de serviços públicos em forma de assistentes virtuais[54]. No que toca a iniciativas privadas, surgiu recentemente o "Alpaca Law"[55], um projeto nacional que possui como objetivo ser um motor de busca para responder a questões que possam ter fundamento direto na letra da lei portuguesa. Por enquanto, trata-se de uma ferramenta de contornos simples e abrangência limitada, mas cremos que poderá ser incentivador para desenvolvimentos futuros no âmbito jurídico nacional.

Além disso, esta realidade também já chegou ao Judiciário. A Inteligência Artificial foi incorporada em vários tribunais brasileiros, incluindo o Supremo Tribunal Federal onde se utiliza o programa "Victor" na organização e avaliação judicial dos processos, utilizado para reconhecer padrões a fim de identificar os temas mais recorrentes nos processos[56], já no Superior Tribunal de Justiça o sistema "Sócrates" promove a análise das peças processuais de recurso, apresenta referências legislativas e cria sugestões de decisão com base nos precedentes[57]. Entre outros exemplos conhecidos em prática, "Dra. Luzia", uma advogada robô utilizada pela Procuradoria-geral do Distrito Federal para analisar processos e elaborar peças no âmbito das execuções fiscais[58]. Salienta-se que o principal objetivo destas ferramentas é sempre auxiliar os magistrados em tarefas rotineiras de modo a facilitar o seu trabalho lógico e mecânico.

Em Portugal, de igual modo, está em curso a implementação de um conjunto de mudanças no âmbito da transformação digital dos tribunais. Para além das melhorias nos sistemas informáticos usados por juízes e procuradores ("Magistraturas" e "MP Codex", respetivamente) que almejam vir a concentrar a tramitação processual num único interface

[52] O serviço foi descontinuado em janeiro de 2021 na sequência de um processo judicial por alegada violação de direitos de autor movida pelos seus concorrentes. *Vide* https://www.rossintelligence.com/features
[53] Para mais exemplos de ferramentas tecnológicas que podem ser alocadas à gestão dos processos judicias cf. ALVES, Paulo F., (2020). "Inteligência Artificial e gestão de grandes processos", pp. 189-197, in ROCHA, Manuel L. & PEREIRA, Rui S. (Coords.), (2020). «Inteligência Artificial & Direito», Coimbra: Almedina.
[54] Por exemplo, o *chatbot* "Sigma" no portal do ePortugal.gov.pt.
[55] https://alpacalaw.com/
[56] Mais sobre os objetivos https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=471331&ori=1
[57] https://www.stj.jus.br/sites/portalp/Paginas/Comunicacao/Noticias/23082020-Revolucao-tecnologica-e-desafios-da-pandemia-marcaram-gestao-do-ministro-Noronha-na-presidencia-do-STJ.aspx
[58] FERRARI, Isabela, *et. al.*, (2018). "Arbitrium ex machina: panorama, riscos e a necessidade de regulação das decisões informadas por algoritmos". *Revista dos Tribunais*. Vol. 995, p. 3.

apoiado por soluções de inteligência artificial, prevê-se também o lançamento de um projeto-piloto no Tribunal Central Administrativo Sul que irá anonimizar decisões judiciais com recurso a um algoritmo de IA.[59] Outra aplicação semelhante refletiu-se no projeto IRIS (Informação, Racionalização, Integração e Sumarização) que será implementado no Supremo Tribunal de Justiça (STJ) e consiste na aplicação de técnicas informáticas e de IA na anonimização e sumarização automática de acórdãos bem como na produção de uma base de dados jurisprudenciais de acesso universal.[60] O desenvolvimento de um interface navegável que capte os elementos-chave dos processos judiciais procurará não só servir de apoio ao processo de decisão, mas também ser um importante instrumento no sentido de facilitar não só a tramitação e divulgação das decisões judiciais mas também a tarefa de análise e compreensão deste tipo de documentos, principalmente para pessoa não técnicas.[61]

3. A JUSTIÇA PREDITIVA

O desenvolvimento tecnológico possibilitou a criação de sistemas computacionais que analisam grandes volumes de dados, identificam padrões e realizam tarefas além da capacidade analítica humana, desafiando o sistema de justiça penal com a sua capacidade de formular juízos de previsão do futuro.[62] A análise preditiva, por si só, não é novidade. Sabemos que é admitida em várias atividades como pontuações de crédito, sugestões de compras online e deteção de fraude bancária. Este tipo de operações auxilia-se em *data mining* e técnicas de *machine learning* para identificar a probabilidade de resultados futuros com base numa grande quantidade de informação recolhida.

Neste âmbito, como terreno de eleição dos instrumentos de IA, surge o conceito de Justiça Preditiva caracterizado como um campo que usa essas tecnologias para prever resultados jurídicos e identificar tendências numa imensa quantidade de dados[63]. Em função

[59] LUSA/DN (2023). "Melhorias informáticas nos tribunais vão poupar deslocações a advogados" https://www.dn.pt/sociedade/melhorias-informaticas-nos-tribunais-vao-poupar-deslocacoes-a-advogados-16571128.html
[60] https://www.inesc-id.pt/projects/PR07005/
[61] *Ibid.*
[62] SOUSA, Susana A. de, (2020b). "Um Direito Penal desafiado pelo desenvolvimento tecnológico: alguns exemplos a partir das neurociências e da inteligência artificial", *Revista da Defensoria Pública da União*, n.º 14, pp. 31-32.
[63] RODRIGUES, (2020a), pp.17-24.

disso, estes modelos podem ser aplicados em diferentes planos, desde antecipar o sentido provável de uma decisão judicial até à probabilidade de um delinquente reincidir.

Já, no seu sentido estrito, a Justiça Preditiva refere-se especificamente a "previsões de resultados de litígios em tribunal"[64] com base em softwares de análise jurisprudencial preditiva. O algoritmo procura fazer correlações de uma grande quantidade de dados encontrados nas decisões judiciais anteriores para exprimir a probabilidade de um juiz decidir em sentido análogo ao decidido em circunstâncias semelhantes.[65]

Porém, a Comissão Europeia ressalva que o termo "justiça preditiva" deve ser evitado porque é ambíguo e falacioso[66]. Não se trata de uma forma de fazer justiça ou julgamentos, apenas de empregar *software* analítico para, por meio de cálculos estatísticos, gerar resultados probabilísticos sobre decisões futuras, sem que as máquinas consigam explicar a relação de causalidade ou fundamentar o raciocínio jurídico que esteve na base das decisões[67]. Esta característica relaciona-se também numa limitação específica da IA: a incapacidade de produzir uma opinião original que a impede de "chegar sozinha a um entendimento próprio sobre" uma mesma situação.[68]

No que toca à possibilidade de prever decisões, já existem disposições legislativas que criminalizam o comportamento de quem utilizar dados sobre os magistrados para calcular o sentido das suas decisões futuras. Falamos da solução legislativa adotada pela França no seguimento da promulgação de uma lei[69] que obriga a disponibilizar online os acervos da jurisprudência produzida no país para consulta de qualquer interessado – uma prática comum na maioria das democracias contemporâneas que reflete o princípio da publicidade processual – com o objetivo de aumentar a transparência sobre o funcionamento dos tribunais.

Numa tentativa de conciliar o acesso público aos dados contidos nos acórdãos redigidos pelos tribunais judiciais franceses com o direito à privacidade das pessoas

---

[64] *Ibid*, p. 17
[65] *Ibid*, p. 32.
[66] CEPEJ (2018) 14. "Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente".
[67] RODRIGUES, (2020a), p. 32.
[68] GUIMARÃES, (2019), p. 1570.
[69] Trata-se da Lei sobre a *Republique Numérique* que impulsionou a transformação digital da França. A *LOI n° 2016-1321 du 7 octobre 2016* pode ser consultada em https://www.legifrance.gouv.fr/

envolvidas, o artigo 33.º da Lei de Reforma do Sistema Judiciário francês[70] modificou o artigo L10-1 do Código de Administração da Justiça e determinou diretrizes para a anonimização da informação relativa à identificação das partes e terceiros mencionados nos acórdãos e proibiu expressamente o tratamento da informação ligada à identidade dos magistrados evitando que esses dados sejam "reutilizados com a finalidade ou efeito de avaliação, análise, comparação ou previsão das suas reais ou supostas práticas profissionais"[71], sob risco de se incorrer em pena de prisão até cinco anos. O objetivo parece ter sido o de evitar que a construção de perfis individualizados pudesse levar a comparações ou previsões sobre os resultados de juízes concretos, bem como pressionar os juízes a decidir de determinada forma.

No entanto, os opositores desta medida reiteram a importância da publicidade das decisões e respetivo *"accountability"*, no sentido de se impor uma prestação de contas jurisdicional[72]. Defende-se que a análise do comportamento judicial traz transparência e legitimação ao funcionamento da justiça, para além de poder constituir uma ferramenta fundamental de pesquisa e análise para todos os operadores judiciários, bem como oferecer um entendimento sobre como as decisões judiciais são formadas[73].

[70] Cf. artigo 33.º da *LOI n° 2019-222 du 23 mars 2019*, que tem como principal objetivo promover a publicidade no Poder Judiciário. Consultar em: https://www.legifrance.gouv.fr/

[71] Cf. https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000038311171/2023-05-17 / Tradução livre.

[72] Contrapondo os argumentos e as posições a favor e contra a previsão das decisões judiciais promovida pela IA, *vide* GUIMARÃES, (2019), *passim.*

[73] Um exemplo de ferramenta de análise jurisprudencial preditiva é a "Jurimetría" espanhola, com uma proposta abrangente que alcança todas as instâncias judiciais e adapta-se a diferentes operadores jurídicos. Mais em: https://jurimetria.laleynext.es/content/Inicio.aspx

IV. A IA COMO FERRAMENTA DE PREVISÃO NO ÂMBITO PENAL

1. NO POLICIAMENTO PREDITIVO

Os algoritmos preditivos (com o auxílio) de IA podem ser aplicados, desde logo, num plano pré-processual, antes de haver uma suspeita ou um crime consumado, naquilo que se veio a apelidar de *(algorithmic) predictive policing*[74], em configurações não muito distantes do ficcionado em "Minority Report", com a exceção de que a previsão dos crimes antes deles acontecerem não cabe a polícias videntes (*"precogs").* O policiamento preditivo do mundo real recorre a "técnicas analíticas – especialmente técnicas quantitativas – para identificar alvos prováveis para intervenção policial e prevenir crimes ou solucionar crimes passados através de previsões estatísticas"[75]. Ora, estas técnicas ao utilizar sistemas de inteligência artificial para realizar a análise preditiva sobre os dados recolhidos, procuram uma maior eficiência na distribuição dos recursos policiais de maneira a evitar ou dissuadir o comportamento criminoso, por exemplo, através do aumento de patrulhamento nos locais e horários em que há maior probabilidade de ocorrências.

Embora as práticas preditivas estejam relacionadas com práticas preventivas e até aplicarmos os dois termos quase indistintamente, do ponto de vista semântico e prático, elas são diferentes. Quando se procura calcular o risco probabilístico de um crime acontecer isso remete-nos para a predição propriamente dita, no sentido em que o que se procura é dispor, de antemão, de uma informação "privilegiada" que indique onde e quando é mais provável o delito acontecer para, de seguida, serem tomadas ações em conformidade com o objetivo traçado, podendo vir a fazer-se uso de técnicas preventivas para neutralizar qualquer possibilidade de ele efetivamente ocorrer. Tanto que a prevenção de crimes, propriamente dita, passa pela incorporação, *in loco*, de estratégias que diminuam a chance de os crimes ocorrerem e, por isso, são um complemento importante da predição de crimes. De igual modo, as técnicas ou ferramentas de deteção de crimes (como sistemas de reconhecimento facial, a monitorização das áreas rurais com drones inteligentes, câmaras de segurança em pontos estratégicos e tecnologias capazes de detetar armas ou drogas, entre outras), ainda

[74] Também pode ser chamado de "Big Data Policing" quando inclui o processamento de uma extensa quantidade de dados aparentemente não relacionados, no sentido de encontrar correlações. BURCHARD, (2021), pp. 179-180

[75] PERRY, Walter L., *et al.*, (2013). *Predictive policing: the role of crime forecasting in law enforcement operations*. RAND Corporation, p. 13. http://www.jstor.org/stable/10.7249/j.ctt4cgdcz

que substancialmente diversas, podem ocorrer ou ser empregues simultaneamente ou em consequência de uma prática de policiamento preditivo ou numa ação de prevenção.

Ou seja, o policiamento preditivo consiste, essencialmente, em recolher dados de variadas fontes, analisá-los de modo que "possa o respetivo resultado ser utilizado pelas polícias, órgãos de investigação criminal e demais entidades responsáveis na prevenção e combate à criminalidade, respondendo assim, de forma mais efetiva e eficaz, à mitigação do risco na prática de futuros delitos criminais"[76]. Mas, uma vez que as previsões são geradas por meio de cálculos estatísticos que, na melhor das hipóteses, produzem estimativas, os resultados serão probabilísticos, não certos.[77] E, ainda que o recurso a ferramentas de estatística e informações geoespaciais no auxílio ao combate ao crime não seja novidade, estes instrumentos têm-se tornado cada vez mais sofisticados e vão ganhando uma maior visibilidade com a ajuda da tecnologia, nomeadamente a utilização de *Big Data*, *ML* e computadores mais potentes[78].

As ferramentas de previsão do risco (*risk acessment tools*) aqui em causa dividem-se em dois tipos principais: técnicas orientadas para determinar o local e a hora com maior risco de ocorrência de crimes – "mapeamento preditivo criminal"[79] –; e técnicas orientadas para a identificação individual (previsão de futuros infratores, i.e., pessoas potencialmente perigosas; identificação de prováveis autores de delitos cometidos no passado; previsão de grupos ou indivíduos suscetíveis de se tornarem vítimas de crimes)[80].

Estas ferramentas, incorporadas em modelos de policiamento, encontram-se muito difundidas nos Estados Unidos. O programa mais conhecido é o PredPol[81], software utilizado na Califórnia para identificar os horários e locais onde os crimes são mais prováveis de ocorrer. Na Europa, também encontramos exemplos similares. Os Países Baixos dispõem do CAS (*"Crime Anticipation System")* para prever futuros focos de crime através de "mapas de calor" que indicam onde e quando o risco de crime é mais elevado[82]. Por sua vez,

---

[76] SILVA, Mário T. da. (2022). "Policiamento preditivo e direitos fundamentais". *Jornal Expresso*. Disponível em https://expresso.pt/opiniao/2022-12-30-Policiamento-preditivo-e-direitos-fundamentais-46fce25a

[77] PERRY, *et al.*, (2013), p. 41.

[78] STRIKWERDA, Litska. (2021). Predictive policing: The risks associated with risk assessment. *The Police Journal: Theory, Practice and Principles, 94*(3), p. 427. https://doi.org/10.1177/0032258X20947749

[79] Quando a informação é reproduzida em "mapa geográfico sob a forma de pontos críticos que são monitorizados em tempo real pelas patrulhas policiais". CEPEJ (2018) 14. "Carta Europeia de Ética (…)", *Op. cit.*

[80] PERRY, *et al.*, (2013)

[81] https://www.predpol.com/about/

[82] STRIKWERDA, (2021), p. 424.

o programa SyRI[83] ("*System Risk Indication")* era utilizado para identificar pessoas físicas ou jurídicas que corriam maior risco de cometer fraude no âmbito laboral, da segurança social e contribuições fiscais[84].

A. DESAFIOS ÉTICO-LEGAIS NO USO DE PROGRAMAS PREDITIVOS

A utilização deste tipo de instrumentos de IA na contenção de crimes apresenta várias limitações, nomeadamente, o facto de não terem em atenção todos os tipos de crimes por ausência de dados suficientes para gerar previsões precisas, sobretudo em crimes menos denunciados e pouco frequentes como o terrorismo ou crimes de colarinho branco[85], também o efeito de "*runaway feedback*" ou "profecias autorrealizáveis" que criam um círculo vicioso na vitimização de grupos ou localidades que já são excessivamente policiadas, bem como a suscetibilidade de qualquer pessoa ser alvo de vigilância[86] ou ficar exposta a possíveis medidas cautelares e de polícia[87], por exemplo, por se encontrar no local que foi identificado como foco provável de um próximo crime ou por ter sido identificada como uma pessoa perigosa.

---

[83] O governo holandês parou de usar SyRI depois do Tribunal Distrital de Haia decidir que este não cumpre com os requisitos de proporcionalidade e transparência necessários e viola o direito à privacidade conforme previsto no art. 8.º da Convenção Europeia de Direitos Humanos, uma vez que acarreta uma intromissão injustificada na vida privada. Mais sobre a decisão *vide* SACHOULIDOU, Athina, (2023). "Going beyond the "common suspects": to be presumed innocent in the era of algorithms, big data and artificial intelligence", pp. 7 e ss.

[84] STRIKWERDA, (2021), p. 423.

[85] SACHOULIDOU, (2023), p. 10.

[86] *Vide*, a este propósito, o artigo 4.º/1 da Estrutura Organizacional da Polícia Judiciária (Decreto-Lei n.º 137/2019, de 13 de setembro) que prevê as competências da PJ em matéria de prevenção e deteção criminal. "a) Promover e realizar ações destinadas a fomentar a prevenção geral e a reduzir o número de vítimas da prática de crimes, motivando os cidadãos a adotarem precauções e a reduzirem os atos e as situações que facilitem ou precipitem a ocorrência de condutas criminosas; 2 – No âmbito da prevenção criminal a PJ procede à deteção e dissuasão de situações conducentes à prática de crimes, nomeadamente através de fiscalização e vigilância de locais suscetíveis de propiciarem a prática de atos ilícitos criminais (…)."

[87] São exemplos de medidas de polícia, segundo o art. 28.º da Lei de Segurança Interna (Lei n.º 53/2008, de 29 de agosto): "a) A identificação de pessoas suspeitas que se encontrem ou circulem em lugar público, aberto ao público ou sujeito a vigilância policial; b) A interdição temporária de acesso e circulação de pessoas e meios de transporte (…)". Contudo, a sua aplicação tem de obedecer ao princípio da necessidade explicitado no art. 30.º do mesmo diploma, i.e., "pelo período de tempo estritamente indispensável para garantir a segurança e a proteção de pessoas e bens e desde que haja indícios fundados de preparação de atividade criminosa ou de perturbação séria ou violenta da ordem pública". Neste sentido, dificilmente se defende que um relatório de risco poderá ser considerado um indício suficientemente fundado, no entanto, quanto mais forte a indicação de "perigo" associada àquele local ou pessoa/grupo concreto, mais pró-ativa poderá ser a atuação dos OPC para encontrar factos que sustentem a suspeita do algoritmo. Cf. art. 55.º/2 e 250.º CPP.

Desde logo, levantam-se dúvidas sobre os parâmetros de risco que possam vir a ser adotados na sua programação, capazes de colidir com garantias de igualdade e não discriminação. A utilização de critérios que se relacionam com características pessoais como a idade, o sexo, etnia, nacionalidade, são parâmetros que fogem ao controlo do delinquente[88] e, por isso, não só apresentam riscos de discriminação, como geram "previsões baseadas em amostras e secundarizam os elementos individualizantes do sujeito"[89]. No mesmo sentido, o uso da variável referente ao código postal de residência é capaz de resultar num "*feedback loop*" que pode perpetuar os padrões existentes de infração[90], "por exemplo, em virtude do sobre-policiamento de determinadas zonas"[91]. Para além dos parâmetros, importa também a quantidade e qualidade dos dados que foram utilizados no "treino" e programação dos respetivos modelos preditivos já que estes irão, inevitavelmente, influenciar a precisão dos seus resultados. O principal receio pode residir na utilização de dados enviesados na alimentação e programação dos modelos preditivos que vão acabar por gerar resultados tendenciosos e "carregados de subjetividade que por vezes replicam preconceitos e discriminações"[92].

Destarte, para além de se levantarem questões sobre a adequação das variáveis analisadas e dos próprios dados utilizados, outras assimetrias surgem quando o visado, numa fase posterior, procura contestar o relatório de risco, principalmente nos casos em que este conduz efetivamente ao exercício de poderes de investigação[93] pelos órgãos de polícia criminal (OPC).

No processo penal, o princípio da presunção da inocência, enquanto princípio estrutural, assume particular importância na fase de investigação durante a qual decorre a maioria das ações de recolha e análise da prova com impacto significativo na futura defesa do suspeito. Porém, quando a suspeita se forma fora do contexto do processo penal, como acontece nos relatórios de risco gerados por meio de policiamento preditivo – "tendo em

[88] SACHOULIDOU, (2023), p. 16.
[89] COSTA, João M. & ABRANTES, António M., (2020). "Os desafios da Inteligência Artificial da Perspetiva Transnacional: A Jurisdição e a Cooperação Judiciária", p. 206, in RODRIGUES, Anabela M. (Coord.), (2020). «A Inteligência Artificial no Direito Penal», Coimbra: Almedina.
[90] SACHOULIDOU, (2023), p. 16
[91] COSTA & ABRANTES, (2020), p. 206.
[92] SALES, Ana D. R., *et al.*, (2021). "Inteligência Artificial e decisão judicial: (im)possibilidade do uso de máquinas no processo de tomada de decisão". *Revista de Processo, Jurisdição e Efetividade da Justiça*. Vol. 7, n.º 1, p. 41
[93] SACHOULIDOU, (2023), p. 10

conta que a geração de um relatório de risco nem sempre, nem automaticamente, resulta na instauração de um processo criminal" – a aplicabilidade do princípio da presunção de inocência é limitada, se não rejeitada, visto que a mera sugestão de suspeita está fora do seu escopo de proteção[94].

Pelo artigo 2.º da Diretiva (UE) 2016/343 do Parlamento Europeu e do Concelho sobre o princípio da presunção de inocência e a proteção dos direitos processuais aplica-se "às pessoas singulares que são suspeitas da prática de um ilícito penal ou que foram constituídas arguidas em processo penal e a todas as fases do processo penal".

Ou seja, ainda que o direito à presunção de inocência, consagrado entre nós no art. 32º/2 da CRP, seja um direito fundamental do arguido e um princípio estruturante do próprio processo penal que se estende até ao trânsito em julgado da condenação, este não tem o mesmo alcance quando o visado é alvo de processamento de dados relevantes para a identificação do seu perfil e, muito menos, quando isso acontece sem qualquer causa prévia ou suspeita individualizada, uma vez que, o âmbito tutelar dos direitos processuais penais "assenta na suspeita da prática de um crime e não no alto risco de vir a cometê-lo no futuro"[95]

Como afirma, Sachoulidou, "um risco ainda não concretizado ou uma intenção abstrata é categoricamente diferente de uma tentativa de crime ou mesmo de atos preparatórios neutros"[96]. No entanto, "a identificação desse risco por meio do policiamento preditivo pode levar os polícias a interpretar cada ação do portador do risco como suspeita, e dar-lhes razões para considerar a interferência com os direitos fundamentais dessa pessoa".[97] Tudo isto, abre portas a novas dimensões de como uma indicação de suspeita pode e deve ser encarada pelos OPC.

Ademais, para fazer previsões, estes sistemas necessitam de uma grande quantidade de dados pelo que a sua recolha, rastreio e análise também pode impactar direitos fundamentais – risco que é agravado quando são empresas privadas as detentoras destes sistemas. O que, por conseguinte, irá também colidir com a capacidade de contestação do relatório de risco se o visado não souber contraditar a forma como o risco foi calculado.

---

[94] *Ibid*, p. 5.
[95] *Ibid*, p. 10.
[96] *Ibid*, *Loc. cit.* Trad. livre.
[97] *Ibid*, *Loc. cit.* Trad. livre.

A recolha e análise de dados, por sua vez, pode levar à vulnerabilização do direito à privacidade e à reserva da intimidade da vida privada e familiar, que é aguçada pela monitorização que as autoridades podem concretizar tendo acesso a diversas fontes de informação (bases de dados, redes sociais, registos nacionais). Neste cenário assume particular relevância o instituto da proteção de dados enquanto pedra angular para limitar a interferência indevida que os dados pessoais podem sofrer.

A este propósito, o Tribunal Constitucional Federal alemão debruçou-se recentemente sobre os riscos que se suscitam na análise e interpretação automatizada de dados durante a prevenção policial, a propósito de uma queixa apresentada pela Sociedade Alemã de Direitos Civis sobre a plataforma de cruzamento de dados "Gotham" em uso no Estado de Hassen, onde é chamada de "HessenDATA", mas também com planos de vir a ser introduzida em Hamburgo. Trata-se de um software de vigilância policial "capaz de pesquisar volumosas quantidades de dados provenientes e recolhidos nos diversos meios de comunicação social e suscetíveis, a final, de permitirem a criação de perfis de suspeitos de crimes, antes mesmo destes ocorrerem"[98]. Ainda que não opere com base em algoritmos de IA, a decisão sobre a constitucionalidade da legislação que admite e regula a sua utilização na vigilância policial, dado os problemas que o tribunal analisou, permite transpor ideias importantes para uma realidade de policiamento reforçada pelas técnicas de IA que suscitem a análise de dados de várias fontes.

O Tribunal constatou que, à partida, os programas de análise poderiam ser justificados à luz da constituição alemã quando verificado o princípio da proporcionalidade tendo em conta o alcance dos poderes em causa. No caso, os sistemas atendem ao legítimo objetivo de aumentar a eficácia na prevenção de atos criminosos graves (como o terrorismo e o crime organizado) "na medida em que permitem descobrir indícios de atos criminosos graves iminentes que, de outra forma, poderiam passar despercebidos nos dados policiais" – apontando que "a análise e interpretação automatizada de dados vai ainda mais longe porque permite o processamento de grandes quantidades de informações complexas"[99].

Porém, um dos pontos levantados pela acusação foi precisamente o acesso que o programa tinha a dados que não estavam relacionados com nenhuma suspeita específica,

[98] SILVA, (2022).
[99] Comunicado de Imprensa n.º 18/2023 do Tribunal Constitucional Federal sobre a Sentença de 16/02/2023. Disponível em: https://www.bundesverfassungsgericht.de/SharedDocs/Pressemitteilungen/EN/2023/bvg23-018.html

cuja análise decorria fora do âmbito de uma investigação legitimadora da interferência, i.e., extravasando o escopo da situação concreta que inicialmente motivou a recolha de dados. Ao procurar por correlações, o sistema acabaria por gerar novas informações, capazes de afetar a vida dos envolvidos, que não seriam acessíveis de outra forma e isso contendia com o fim para os quais aqueles dados foram inicialmente recolhidos. O tribunal estabeleceu que o limite constitucionalmente exigido aqui para admitir a interferência nos direitos fundamentais dos visados devia ser o da existência de "um perigo suficientemente identificável"[100]. Notando, porém, que a base legal existente no momento em Hasse, ao apresentar uma redação particularmente ampla dos poderes concedidos, tanto em termos de dados utilizados como dos métodos envolvidos, não atendia a essa critério[101] e permitia a "realização de cruzamentos de dados em abstrato sem vinculação a um perigo justificadamente concreto"[102]. Outra preocupação do tribunal apontou o perigo para a formação de perfis de pessoas e grupos com apenas um *click*, decorrente do potencial do software conforme os contornos definidos naquela lei para a modalidade de prevenção de infrações penais graves[103], o que poderia sujeitar muitas pessoas legalmente inocentes, mas identificadas erroneamente como suspeitas, a outras medidas policiais[104].

Posto isto, o tribunal declarou que as legislações utilizadas em Hesse e Hamburgo relativas à análise e interpretação automatizada de dados para a prevenção de atos criminosos, nas suas redações atuais, eram inconstitucionais por violarem o "direito à autodeterminação informacional"[105]. Também foram estabelecidos diversos critérios ao longo da decisão que o legislador deveria seguir para balizar a intromissão nos direitos fundamentais decorrente do tratamento automatizado de dados principalmente nos casos em que ocorre uma mudança da finalidade no seu tratamento pelo Estado. Ordenando, portanto, que a legislação fosse adaptada para definir com clareza as condições sobre as quais as autoridades policiais poderiam processar, de forma legal, os dados recolhidos, restringindo necessariamente os motivo de intervenção permitidos.

---

[100] *Ibid*, Trad. livre.
[101] *Ibid.*
[102] CAMPOS, Ricardo (2023). "Autodeterminação informacional 4.0 e o tratamento de dados pelo Estado". *Revista Consultor Jurídico*. Disponível em: https://www.conjur.com.br/2023-fev-28/direito-digital-tratamento-dados-estado-limites-constitucionais#_ftn11
[103] *Ibid.*
[104] Comunicado de Imprensa n.º 18/2023 (…), *Op. Cit.*
[105] *Ibid.*

Esta decisão é ainda relevante na medida em que convoca o direito à autodeterminação informacional[106] para ambientes automatizados alimentados por uma imensa quantidade de dados – uma configuração que exige um reforço da proteção dos dados pessoais que ficam ainda mais expostos a impactos negativos.

Transpondo a problemática para o nível nacional, verificamos que essa proteção encontra-se devidamente acautelada pela sinergia entre a Lei 58.º/2019 (assegura a execução do Regime Geral de Proteção de Dados na ordem jurídica interna) e a Lei n.º 59/2019 (transpõe a Diretiva (EU) 2016/680) que se aplica, precisamente, ao tratamento de dados pessoais por meios total ou parcialmente automatizados "para efeitos de prevenção, deteção, investigação ou repressão de infrações penais ou de execução de sanções penais". Daqui resulta que, entre nós, é permitido o tratamento dos dados pessoais para finalidades diferentes daquelas para as quais os dados pessoais foram recolhidos inicialmente desde que essas finalidades ainda se enquadrem nos fins da lei e a lei assim o autorize (art. 7.º da Lei n.º 59/2019), mas, por outro lado, a lei portuguesa proíbe, independentemente do consentimento do titular, a tomada de decisões "exclusivamente com base no tratamento automatizado, incluindo a definição de perfis, que produzam efeitos adversos na esfera jurídica do titular dos dados ou que o afetem de forma significativa, exceto quando autorizadas por lei, desde que seja previsto o direito de o titular dos dados obter a intervenção humana do responsável pelo tratamento" (art. 11.º). No mesmo sentido, quando o tratamento de dados seja processado por IA, o art. 9.º da Carta Portuguesa de Direitos Humanos na Era Digital[107] acrescenta que essa aplicação "deve ser orientada pelo respeito dos direitos fundamentais, garantindo um justo equilíbrio entre os princípios da explicabilidade, da segurança, da transparência e da responsabilidade, que atenda às circunstâncias de cada caso concreto e estabeleça processos destinados a evitar quaisquer preconceitos e formas de discriminação" e prevê, igualmente, "que as decisões com impacto significativo na esfera dos destinatários que sejam tomadas mediante o uso de algoritmos devem ser comunicadas aos interessados, sendo suscetíveis de recurso e auditáveis, nos termos previstos na lei".

[106] Reconhecida pela primeira vez numa decisão do Tribunal Constitucional Federal alemão, em 1983, como o direito de cada pessoa para decidir sobre "a divulgação e o uso dos seus dados pessoais". Entre nós este direito pode ser deduzido do art. 35.º CRP. Cf. PEREIRA, Alexandre L. D., (2019). "O Responsável pelo Tratamento de Dados Segundo o RGDP". *Revista de Direito e Tecnologia*, Vol. 1, n.º 2, pp. 148-149.
[107] Lei n.º 27/2021, de 17 de Maio.

2. NA TOMADA DE DECISÃO DURANTE O PROCESSO

Durante o processo, agora com o foco numa suspeita já estabelecida, o recurso a algoritmos de avaliação de risco *(risk assessment algorithms)* que analisam uma variedade de fatores relacionados com a perigosidade criminal do agente – baseados em variáveis atinentes à pessoa, tanto respeitantes à sua história criminal como às suas características sociodemográficas –, e a aplicação, no geral, da IA à política punitiva, prometem a tomada de decisões[108] de forma mais alinhada com o objetivo de alcançar a desejada racionalidade no momento da punição, assumindo que "os algoritmos preditivos sumariam a informação relevante de uma maneira mais eficiente do que o cérebro humano"[109]. Localizamo-nos, agora, num plano mais sensível da aplicação da justiça, onde as consequências, fruto de decisões automatizadas, ditas inteligentes, serão, certamente, mais gravosas e impactantes de direitos fundamentais.

Ainda que normalmente "se ressalve a aplicação para o mais notável ato do processo, a sentença"[110] – onde assume particular importância a problemática da reincidência, estreitamente relacionada com a operação de escolha da pena e a operação da determinação da medida concreta da pena[111] – é possível que se admita a utilização da IA "para atos processuais diversos"[112].

No plano processual penal, a atividade jurisdicional envolve vários momentos que exigem ponderação e decisão, muitas vezes discricionária, e atos que podem envolver a realização de juízos prognósticos, incluindo avaliações de risco, que a IA embutida nos algoritmos preditivos pode ajudar a calcular[113]. Desta forma, a justiça criminal torna-se apelativa para integrar as tecnologias de IA com o objetivo de facilitar as decisões de quem investiga, acusa e julga, durante os vários momentos da realização da justiça penal, assumindo "especial relevância nos casos em que a decisão do julgador deva ser adotada segundo um juízo de probabilidade ou de proporcionalidade"[114]. Conforme a fase judicial

[108] RODRIGUES, Anabela M., (2020b). «Medida da pena de prisão – desafios na era da inteligência artificial», Revista de Legislação e de Jurisprudência, Vol. 149, n.º 4021, pp. 264-265.
[109] *Ibid,* p. 261.
[110] PEDRINA, Gustavo M. L., (2019). "Consequências e perspectivas da aplicação de inteligência artificial a casos penais", Revista Brasileira de Direito Processual Penal, Porto Alegre, set.-dez., Vol. 5, n.º. 3, p. 1597.
[111] SANTOS, Hugo L. dos, (2022). "Inteligência Artificial e Processo Penal". Nova Causa, pp. 64-65.
[112] PEDRINA, (2019), p. 1597.
[113] COSTA & ABRANTES, (2020), p. 202
[114] GONÇALVES, (2022), p. 285.

em causa, o seu responsável – órgão policial, Ministério Público, ou juiz – é instado a fazer um juízo de previsão sobre o comportamento futuro do investigado, do acusado ou do condenado[115].

Veja-se o caso da aplicação de uma medida de coação na fase de inquérito para fazer face a exigências cautelares como o risco de fuga, continuação da atividade criminosa ou adulteração provas (artigos 193.º/1 e 204.º/1 CPP)[116]; também na fase de julgamento, o momento da escolha (entre uma pena principal não privativa da liberdade e uma privativa, ou de uma pena de substituição[117] em detrimento de uma pena principal, dependendo da realização adequada e suficiente das finalidades da punição[118]) e da determinação da medida concreta da pena[119], baseada no juízo de índole prognóstica sobre o risco de reincidência do agente; de igual modo, a decisão de aplicar uma medida de segurança que envolverá necessariamente um juízo de prognose sobre a perigosidade do agente considerado inimputável em razão de anomalia psíquica (artigos 40.º/3 e 91.º/1 CP). O objetivo é prever o comportamento futuro do arguido, na tentativa de aumentar a segurança e uma aplicação mais ajustada às necessidades de prevenção.

"O processo de racionalização[120] do *sentencing* (quer no que tange à operação da escolha da pena, quer no que respeita à determinação da medida concreta da pena) assenta" numa tentativa de reduzir o diferencial de subjetividade que se lhe encontra associado.[121]

---

[115] SOUSA, Susana A. de, (2020). "Um direito penal desafiado pelo desenvolvimento tecnológico: alguns exemplos a partir das neurociências e da inteligência artificial", *Revista da Defensoria Pública da União*, p. 32.

[116] Este juízo de prognose será, no entanto, diferente daquele que se faz no momento da determinação da medida da pena, devendo ser analisado estritamente à luz das exigências processuais de natureza cautelar. Cf. ANTUNES, Maria J. (2022). «Direito Processual Penal», Coimbra: Almedina, 4ª ed. pp. 167-168.

[117] No caso da suspensão da execução da pena "está sempre uma prognose social favorável ao agente, baseada num risco prudencial. Porém, o juízo de prognose que o tribunal faz não tem carácter discricionário e, muito menos, arbitrário. O tribunal ao decretar a medida terá de refletir sobre a personalidade do agente, sobre as suas condições de vida, a sua conduta *ante et post crimen* e sobre o circunstancialismo envolvente da infração." Cf. Ac. do STJ de 9/01/2002, Processo n.º 3026/01 – 3.ª secção, https://www.stj.pt/wp-content/uploads/2018/01/criminal2002.pdf

[118] São as necessidades de prevenção que vão justificar a opção da pena, como resulta dos critérios estabelecidos nos arts. 40.º/1 e 70.º CP, não existindo aqui qualquer finalidade de compensação da culpa, uma vez que esta, constituindo o limite da pena (art. 40.º/2 CP), apenas funciona ao nível da determinação da medida concreta.

[119] Feita em função da culpa do agente e das exigências de prevenção reclamadas pelo caso concreto (art. 71.º/1 CP). Sobre a dificuldade de vincular uma grandeza numérica da pena a qualquer critério normativo que a submeta a uma única e correta correspondência entre a gravidade do caso e a sua expressão aritmética, *vide* RODRIGUES, (2020b).

[120] Numa lógica de cunho tecnocrático e economicista subjacente a uma análise de custo-benefício e utilidade-eficiência na alocação e otimização dos recursos escassos da administração da justiça penal. Cf. SANTOS, (2022); RODRIGUES, (2020b), p. 263.

[121] SANTOS, (2022) p. 70.

Esta finalidade encontra maior expressão nos sistemas de *Common Law* que procuram cada vez mais introduzir critérios quantitativos no processo decisório com vista a alcançar uma maior racionalidade e eficácia na avaliação de risco – cenário para o qual "contribui também um sistema penal baseado cada vez mais na *guilty plea* e na *plea bargaining*, com efeito burocratizante na discussão do caso"[122] – o que potencia a utilização de instrumentos preditivos, face a uma justiça europeia continental que continua a privilegiar a "fase da discussão da culpa, com intensa intervenção humana"[123] e apreciação do caso concreto ao encargo do julgador, o que resulta, por sua vez, num sistema menos permeável à entrada de sistemas preditivos.

Já no plano da execução das sanções, o momento da concessão da liberdade condicional também depende de um juízo de prognose por parte do julgador sobre como será o comportamento do recluso no que respeita à reiteração criminosa no futuro e que, segundo o art. 61.º CP, deve ser "fundadamente de esperar (…) que o condenado, uma vez em liberdade, conduzirá a sua vida de modo socialmente responsável, sem cometer crimes". Por último, Costa e Abrantes alargam ainda a aplicação da IA à monitorização do cumprimento de certas formas de execução da pena de prisão, *e.g.*, regime de permanência na habitação.[124]

A CEPEJ constata que, apesar dos testes e estudos que vão sendo feitos para explorar o potencial destas aplicações, atualmente, os juízes dos países membros do Conselho da Europa não utilizam na prática diária nenhum software preditivo[125]. Aquilo que se apresenta como uma ideia "estranha" e particularmente rara na Europa, vai, no entanto, abundando na literatura estadunidense, sobretudo devido às diferenças dos modelos de justiça penal praticados. A experiência norte-americana já conta com mais de sessenta instrumentos de avaliação do risco atualmente em uso nos vários estados[126]. Recorre-se aos *risk assessment tools*, movidos a algoritmos, em tomadas de decisão nas diversas fases do sistema de justiça penal, nomeadamente para sustentar decisões sobre *bail*, *sentencing* e a

[122] RODRIGUES, (2020a), p. 20.
[123] *Ibid, Loc. cit.*
[124] COSTA & ABRANTES, (2020), p. 205.
[125] "As iniciativa para o desenvolvimento destes instrumentos provém em grande parte do sector privado, cuja clientela até agora tem sido maioritariamente constituída por companhias de seguros, advogados e serviços jurídicos que pretendem reduzir a insegurança jurídica e a imprevisibilidade das decisões judiciais". CEPEJ (2018) 14. "Carta Europeia de Ética (…)", *Op. cit.*
[126] RODRIGUES, (2020b), p. 264.

liberdade condicional de um recluso com o objetivo de "introduzir eficácia e racionalidade na aplicação da punição"[127].

No entanto, ao oferecer de maneira autónoma a melhor resposta estatística para a questão colocada, "o que quase todas as aplicações existentes fazem é computação estatística" – técnica que faz parte das teorias de *machine learning*[128]. A tendência remonta ao fenómeno chamado de "*evidence-based sentencing*" (EBS) na formulação técnico-legal anglo-americana, que em tradução livre significa "condenação baseada em dados"[129] e está associada à prática que envolve a análise de pesquisas científicas e estatísticas sobre o impacto que diferentes sentenças e intervenções têm na reincidência criminal, reabilitação, segurança pública e outros aspetos sociais[130]. Ocorre que apesar do nome, "*evidence*", neste contexto, refere-se, não à prova do caso concreto, mas, aos dados objetivos relativos ao sujeito, como género, endereço e histórico criminal que assume um peso considerável na equação[131]. O objetivo é ajudar os juízes a tomarem decisões de sentença mais informadas e promover a eficácia do sistema de justiça criminal e a redução da criminalidade no alcance dos objetivos da prevenção. Acabando, no entanto, por desligar-se a punição "da gravidade do facto praticado para passar a ser proporcionada ao risco que o indivíduo representa para a sociedade"[132].

O uso de IA durante os julgamentos criminais já foi contestado em vários tribunais dos Estados Unidos. Um dos exemplos mais mediáticos e controversos de determinação da pena pela máquina ocorreu no caso *Loomis vs. Wisconsin*, nos EUA, onde o programa de inteligência artificial COMPAS (*Correction Offender Management Profiling for Alternative Sanctions*[133]), classificou o réu como um indivíduo de alto risco (leia-se, com uma alta

---

[127] *Ibid, Loc. cit.*
[128] PEDRINA, (2019), pp. 1592
[129] O termo não é utilizado no contexto jurídico português, mas a expressão pode ser encontrada em RODRIGUES, (2020a), p. 48.
[130] PEDRINA, (2019), pp. 1594; STARR, Sonja B., (2013). "Evidence-Based Sentencing and the Scientific Rationalization of Discrimination". *Stanford Law Review*, Forthcoming, U of Michigan Law & Econ Research Paper n.º 13-014. https://ssrn.com/abstract=2318940
[131] *Ibid.*
[132] RODRIGUES, (2020a), p. 48.
[133] Produzido pela empresa *Equivant*, o COMPAS é um programa de avaliação de risco usado em vários estados americanos. O programa COMPAS usa um algoritmo de IA para calcular o risco de reincidência de um réu com base em informações por si prestadas num questionário de 137 tópicos (incluindo idade, género, emprego, educação, entre outras variáveis) e o seu histórico criminal. Tem como objetivo auxiliar os juízes na tomada de decisões sobre a aplicação de uma pena de prisão, a liberdade condicional ou a fiança antes do julgamento. Segundo os produtores, o programa analisa um conjunto de dados de uma amostra de uma população relevante – informações sobre grupos de indivíduos em situações análogas que já passaram pelo sistema de justiça

probabilidade de voltar a cometer um crime), levando o juiz a aplicar – baseado, em parte, nessa avaliação – uma pena efetiva de seis anos de prisão acompanhada de cinco anos de supervisão estendida[134]. Loomis recorreu até a Suprema Corte de Wisconcin questionando os critérios que levaram o algoritmo a recomendar a sua classificação, invocando que isso violou o seu direito a um *fair trial*, bem como o direito a uma sentença individualizada e de ser julgado com base em informações precisas – a Suprema Corte estadunidense rejeitou o recurso, considerando que a decisão teria sido a mesma independentemente do uso da avaliação de risco algorítmica pelo tribunal e essa aplicação pelo tribunal de primeira instância não violou o seu direito a um processo equitativo *(due process)*, pese embora a metodologia usada para realizar a avaliação não ter sido divulgada nem ao tribunal nem ao réu uma vez que se tratava de segredo comercial (*trade secret*)[135]. O programa COMPAS, bem como outras ferramentas de previsão do risco[136], foi alvo de duras críticas que argumentavam que este pode ser tendencioso e discriminar certos grupos, como afro-americanos e latinos[137].

Esta jurisprudência deixou, contudo, algumas recomendações práticas que os tribunais devem observar para evitar potenciais violações do *due process* no caso de se usarem instrumentos como o COMPAS[138]. Mais especificamente: (1) as avaliações de risco não podem ser usadas como fator determinante na sentença; (2) devem ser explicados os fatores, para além da avaliação de risco, que contribuíram para a fundamentação da sentença; (3) as avaliações de risco não devem ser usadas para determinar se alguém será condenado

---

criminal – para identificar os fatores considerados mais importantes para desenvolver uma avaliação do risco de reincidência e, com base nessa experiência, estima um valor para o visado, sem atender, no entanto, as circunstâncias concretas daquele réu. Cf. CARIA, Rui, (2020). "O caso State v. Loomis – a pessoa e a máquina na decisão judicial", pp. 245-265, in RODRIGUES, Anabela M. (Coord.), (2020). «A Inteligência Artificial no Direito Penal», Coimbra: Almedina.

[134] State v. Loomis. 130 *Harvard Law Review* 1530 (2017, march). https://harvardlawreview.org/print/vol-130/state-v-loomis/

[135] *Ibid.*

[136] O sistema HART (*Harm Assessment Risk Tool*) é o exemplo de ferramenta preditiva de risco utilizada na Inglaterra. Com base em 34 indicadores, tem como objetivo determinar o risco de reincidência durante os próximos dois anos de pessoas que foram detidas por cometerem crimes e avaliar a possibilidade de estas serem submetidas a um programa de reabilitação chamado "*CheckPoint"*. SACHOULIDOU, (2023), p. 16

[137] A crítica vertida numa pesquisa realizada pela organização *"ProPublica"* baseia-se na preocupação de que as pontuações do COMPAS, ao serem influenciadas por fatores como a raça, podem gerar resultados discriminatórios. O estudo concluiu que os cidadãos negros eram mais vezes avaliados como de alto risco em comparação com arguidos caucasianos, em relação aos mesmos delitos ou de gravidade menor. Cf. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

[138] SACHOULIDOU, (2023), p. 18.

ou quão severa será a sua sentença; e (4) os tribunais que implantam o COMPAS devem receber aconselhamento por escrito *(warning label)* sobre as suas limitações.[139]

A. GARANTIAS E DIREITOS PROCESSUAIS

O caso *Loomis* foi importante para desencadear o debate sobre a possibilidade de utilizar algoritmos com suporte de IA para auxiliar na tomada de decisões em questões criminais perante os tribunais europeus[140]. À primeira vista, facilmente identificaríamos os benefícios de tal introdução no sistema judiciário português, já que "os argumentos a favor concentram-se no direito a uma audiência dentro de um prazo razoável e o direito a tal audiência por um tribunal independente e imparcial como elementos centrais de um processo equitativo."[141]

O emprego de tecnologias digitais e ferramentas de IA, munidos de uma grande capacidade de análise algorítmica, durante a tramitação processual no auxílio ao juiz no manuseamento e análise do processo, bem como nas tomadas de decisão, não necessariamente limitadas à condenação, pode, de facto, contribuir para diminuir a morosidade[142] dos atos processuais e assegurar "os valores do processo inerentes à efetividade e celeridade do exercício da acção penal, nomeadamente conservar as provas, garantir as finalidades cautelares das medidas de coação ou evitar a prescrição do procedimento criminal"[143] – o que reveste particular utilidade nos "megaprocessos", caracterizados pela complexidade das matérias que abordam e uma esmagadora quantidade de prova recolhida, bem como nos processos urgentes que requerem diligências atempadas na prevenção de danos potencialmente irreparáveis.

Invoca-se, a propósito, o direito a uma decisão em prazo razoável integrante da garantia constitucional de "acesso ao direito e tutela jurisdicional efetiva" (art. 20.º/4 CRP): "todos têm direito a que uma causa em que intervenham seja objeto de decisão em prazo

---

[139] *Ibid.*
[140] SACHOULIDOU, (2023), pp. 18 e ss.
[141] *Ibid.*
[142] Segundo o Painel de Avaliação da Justiça na UE de 2023, Portugal é o quinto país da UE com maior duração dos processos nos tribunais, sendo o mais lento no que toca às decisões administrativas das autoridades de proteção dos consumidores. Notícia em https://www.tsf.pt/portugal/sociedade/portugal-e-o-5-pais-da-ue-com-maior-duracao-dos-processos-nos-tribunais-16496519.html
[143] Como sufragado pelo Ac. do Tribunal da Relação do Porto, 5/04/2017, Processo n.º 189/12.6TELSB-AB.P1, que aponta a efetividade e celeridade no exercício da ação penal como sendo também uma finalidade do processo. Disponível para consulta em http://www.dgsi.pt.

razoável e mediante processo equitativo". Também o n.º 1 do art. 6.º da Convenção Europeia dos Direitos do Homem[144], com a epígrafe "Direito a um processo equitativo", manifesta-se nesse sentido: *"Qualquer pessoa tem direito a que a sua causa seja examinada, equitativa e publicamente, num prazo razoável por um tribunal independente e imparcial, estabelecido pela lei (...)."*

"É por referência ao direito ao acesso à justiça que se afigura crucial considerar-se que a aplicação da nova dimensão tecnológica assume um especial contributo na otimização das tarefas diárias do juiz"[145] que procura providenciar pelo desenvolvimento célere do processo – não esquecendo que, "na apreciação da razoabilidade da duração de um processo há que atender à complexidade jurídica do caso e às diligências cuja realização do mesmo reclama"[146]. Isto relaciona-se com a competência (poder-dever) do juiz presidente, enquanto *dominus* do processo, de exercer uma eficaz e eficiente gestão processual adequada à justa composição do litígio em prazo razoável (art. 6.º CPC, aplicável subsidiariamente ao processo penal por força do art. 4º CPP).

Engelmann e Fröhlich apontam para o perigo de utilizar a IA para além das questões burocráticas, convencidos de que sustentar a legitimação da automaticidade das decisões na garantia da efetivação da duração do prazo razoável é privilegiar a quantidade em detrimento da qualidade na produção judicial[147].

No que toca às exigências materiais de independência e imparcialidade do tribunal[148], espera-se que a produção judicial seja "livre" de interferências externas[149] e critérios subjetivos, sustentada e decidida exclusivamente naquilo que é prescrito pelo

---

[144] Outra garantia processual vertida no n.º 3, al. b), do mesmo artigo – bem como no art. 32.º CRP – refere-se ao direito do acusado de dispor de tempo e meios necessários para preparar a sua defesa. Neste sentido, a favor do recurso à IA para otimizar a prestação de serviços por parte dos advogados de defesa, em referência ao Caso Marquês, cf. MEDEIROS, (2020); GONÇALVES, (2020), todos in ROCHA, Manuel L. & PEREIRA, Rui S. (Coords.), (2020). «Inteligência Artificial & Direito», Coimbra: Almedina.

[145] PINHOL, Ana M. M. F., (2022). "A inteligência artificial e a justiça encontros e desencontros", Lisboa: Centro de Estudos Judiciários, p. 45. Disponível em: https://cej.justica.gov.pt/LinkClick.aspx?fileticket=ttrgZIvgv7E%3D&portalid=30

[146] *Ibid,* p. 44.

[147] ENGELMANN, Wilson & FRÖHLICH, Afonso V. K., (2020). "Inteligência Artificial Aplicada À Decisão Judicial: O Papel dos Algoritmos no Processo de Tomada de Decisão". *Revista Jurídica* (FURB), Vol. 24, N.º 54.

[148] O princípio da independência judicial encontra-se plasmado no art. 203.º CRP que, em conjunto com o art. 32.º/5 CRP referente à estrutura acusatória do processo penal, sustenta também o princípio da imparcialidade.

[149] Enquanto reflexo do corolário da independência externa dos juízes, numa concretização do princípio da separação de poderes. Já a independência interna significa que os tribunais são independentes uns dos outros e entre si.

sistema normativo, de acordo com o direito vigente, de forma distante e desinteressada das próprias vontades e inclinações do juiz. Assim, o direito de cada cidadão em ter a sua causa decidida por juízes imparciais e independentes pretende garantir que as decisões são alcançadas através de um processo que respeite os direitos e liberdades fundamentais.

A questão que se pode colocar aqui é se o auxílio de algoritmos preditivos reforça ou aniquila o voto de confiança que a comunidade deposita nos juízes na expetativa destes decidirem de forma isenta e neutral, visto que as críticas que em larga medida se convocam contra uma configuração algorítmica, são precisamente, os resultados enviesados que estes podem gerar e a opacidade do método pelo qual produzem esses *outputs.*

Porém, do lado mais otimista, os adeptos desta introdução, acreditam ser este um caminho possível para contornar os vieses, preconceitos e erros *(cognitive biases)* inerentes aos magistrados enquanto seres humanos, afastando julgamentos arbitrários e garantindo aos réus oportunidades iguais perante os tribunais, já que todos seriam submetidos aos mesmos mecanismos de avaliação e decisão. Em contrapartida, ao retirar o poder da arbitrariedade ao juiz, o algoritmo incorre no risco de suprimir também a necessária discricionariedade das decisões judiciais, onde entram o pensamento intuitivo, as valorações[150], experiências pessoais e a empatia humana, que assumem um valor inestimável na apreciação das circunstâncias únicas do contexto de cada caso concreto. Será pertinente adaptar aqui a essência da velha máxima "*summum ius, summa iniuria"* que neste sentido se podia insurgir contra uma justiça "completamente previsível em aplicação automática".[151]

Por outro lado, contra essa contingência também parecem contribuir certas especificidades da atividade jurídica que, no estado atual da técnica, muito dificilmente se concretizam com recurso a mecanismos automáticos de IA, como acontece com a interpretação de normas, de princípios jurídicos e conceitos indeterminados, que, por enquanto, estão apenas ao alcance da capacidade humana.

De igual modo, a impossibilidade por parte de sistemas de IA de dar resposta às lacunas, resolver situações em que exista colisão de direitos e "análise da prova testemunhal, e, por conseguinte, a impossibilidade sensitiva imperatória na análise do caso concreto" parecem afastar, no entendimento de Patrícia Borges, que subscrevemos, na substituição

[150] RODRIGUES, (2020a), p. 26.
[151] PINHOL, (2022), p. 50.

total do juiz humano por um "juiz-robô".[152] Contra esta possibilidade – e em forma de um breve parênteses –, entre outras objeções possíveis, um impedimento legal sobressai de imediato: a colisão com o princípio do "juiz natural"[153] espelhado no art. 32.º/9 CRP. Ao ditar que "nenhuma causa pode ser subtraída ao tribunal cuja competência esteja fixada em lei anterior", o princípio obriga a que a retirada da causa ao juiz competente tenha sempre base legal. Assim sendo, qualquer tentativa de delegar nos sistemas de IA a decisão da causa penal, sem uma lei que admitisse expressamente essa substituição, numa espécie de cedência das funções do julgador, seria atentatória das garantias plenas de uma justiça independente e imparcial que depositam naquele juiz concreto a expetativa da decisão justa, mas cuja causa, desta forma, é lhe retirada. A questão sobre a viabilidade de inclusão de uma máquina dotada de IA para "produzir" decisões finais que seriam válidas de acordo com as configurações de um certo sistema jurídico, o tal "juiz robô", é uma questão mais profunda de cariz jurídico-filosófico que não iremos aprofundar aqui[154].

Retomando a análise dos inconvenientes que a automatização da decisão levanta, e assumindo que, ainda que não se trate de sistemas ou expedientes intrinsecamente injustos, o seu emprego pelas autoridades judiciárias, pode promover uma desproporção de meios entre a acusação e a defesa[155] no sentido em que a opacidade do funcionamento do sistema dificulta, de forma ilegítima, o exercício do direito de defesa na vertente do contraditório, bem como do direito ao recurso, precisamente pela razão de que não se consegue contestar eficazmente o que não se conhece.

Isto relaciona-se com as características da insindicabilidade do algoritmo e falta de transparência dos métodos utilizados, que acaba por limitar a compreensão dos trâmites da tomada de decisão pelo sistema. Consequentemente, persiste a dificuldade em perceber e descrever como se processam os cálculos e se chega a um certo resultado ou qual o valor que se atribui a cada indicador de avaliação face ao seu peso no veredicto final. Este impedimento consubstancia o problema da "caixa negra" *(black box problem)* característico

[152] BORGES, Patrícia S., (2022). «A utilização de inteligência artificial na justiça aos olhos da proposta de regulamento da União Europeia», pp. 287-304.
[153] *Ibid.*
[154] Sobre o assunto, e em especial o confronto entre o raciocínio jurídico e a IA, *vide* PEREIRA, Alexandre L. D., (2020). "Inteligência Artificial na Decisão Jurisprudencial?". *JURISMAT*, Portimão, n.º 12, pp. 73-92.
[155] COSTA & ABRANTES, (2020), p. 206.

da opacidade das tecnologias de IA que culmina numa ausência de explicabilidade[156]. Muitas vezes, a razão do impedimento, coberta pelo prorrogativa de proteção da propriedade intelectual, resulta do facto de serem empresas privadas as detentoras do "segredo" do *software*. Tudo o suprarreferido coloca em causa, de forma generalizada, o direito a um processo justo e equitativo.

À medida que os algoritmos de IA evoluem e são utilizados em domínios críticos, maior é a exigência sobre a necessidade de se criarem condições que permitam aos utilizadores ou destinatários "compreender o que levou o sistema a tomar determinadas decisões algorítmicas"[157]. A explicabilidade dos cálculos relaciona-se diretamente com a exigência de transparência. A concretização da transparência promove a confiança na tecnologia que, por sua vez, é uma condição prévia para a sua aceitação[158].

Outro ponto aguçado nesta discussão, apontado por alguns autores, remete ao perigo de se cair na tendência de desligar a punição da gravidade do facto e associar a avaliação quantitativa de risco individual a uma espécie de "*scoring*" pessoal, neste caso, sobre o risco de perigosidade ou reincidência que o individuo representa para a sociedade. Ao permitir que se crie um "perfil algorítmico" do suspeito ou arguido faz-se "desaparecer o facto praticado e a discricionariedade é erradicada e substituída pelo determinismo"[159]. Corre-se o risco de o delinquente ficar adstrito a um futuro determinado por um cálculo em função das suas características passadas avaliadas por instrumentos de IA que lhe prescrevam um mau *scoring* que irá condicionar as decisões sobre a sua condenação[160] e colocar em causa a responsabilidade pela sua autodeterminação. Ainda que a IA consiga acertar facilmente na previsão de eventos que ocorrem com frequência e em grande número,

---

[156] Ressalva-se, porém, que nem toda a IA será uma tecnologia de "caixa negra". Já começam a surgir também modelos de "white-box" que procuram ser interpretáveis e consubstanciam aquilo que se chama *"eXplainable Artificial Intelligence"* (XAI) – um conjunto de técnicas e métodos que convertem os algoritmos de "caixa preta" em algoritmos de "caixa branca". Cf. HUSSAIN, Fatima, *et. al.*, (2021). "Explainable Artificial Intelligence (XAI): An Engineering Perspective". https://doi.org/10.48550/arXiv.2101.03613

[157] COM (2020) 64 final. Bruxelas, 19.2.2020.

[158] COM (2020) 65 final. Bruxelas, 19.02.2020 – Livro Branco sobre a inteligência artificial. De igual modo, o relatório do Grupo de Peritos de Alto Nível sobre a IA, (2019). "Orientações éticas para uma IA de confiança" enumera um conjunto de 7 requisitos essenciais aos quais os sistemas de IA devem atender para serem considerados de confiança.

[159] "Perspetiva-se o comportamento humano no enquadramento preditivo de um computador e de acordo com uma nova normatividade algorítmica (feita de correlações), que pesa, não apenas a favor da renúncia à discricionariedade do julgar, como, além disso, da progressiva desvalorização do momento do julgamento". RODRIGUES, (2020a), p. 48.

[160] *Ibid,* p. 47.

a confiabilidade na natureza probabilística da previsão de risco diminui drasticamente quando se faz uma previsão de nível individual.

Mesmo que o futuro pudesse ser conhecido com certeza, punir alguém com base em atos futuros pode ser fundamentalmente injusto e contraditório da própria teoria da punição porque "pune-se o réu por quem ele é (e o que, portanto, espera-se que ele faça no futuro), em vez do que ele fez"[161].

B. RESPONSABILIDADE HUMANA NAS DECISÕES AUTOMATIZADAS

No que toca à inclusão dos sistemas preditivos no âmbito da justiça para auxílio em juízos de prognose pontuais – ainda que acessórios à decisão principal –, onde o julgador vem apenas colher da vasta capacidade de análise da IA para tomar decisões mais ponderadas, releva, a possibilidade de o juiz poder discordar do resultado gerado e oferecido como certo pelos algoritmos. Para tanto, importa saber que tipo de resultados a máquina produz e como deveria o juiz interpretar os dados dali resultantes, nomeadamente quando consubstanciem uma avaliação do risco.

Considerando que os resultados "produzidos" pelo sistema preditivo possam constituir elementos relevantes para interferir na determinação da sanção aplicável[162] – por exemplo, na vertente das exigências de prevenção –, o relatório de risco que surge, quando suficientemente claro e autoexplicativo, deve ser considerado um meio de prova e, como tal, sujeito ao princípio da livre apreciação da prova nos termos do art. 127.º CPP.

O mesmo sentido foi dado no caso *Loomis* ao entender-se que "o COMPAS é uma ferramenta de auxílio à decisão e como tal será uma prova como outra qualquer, sujeita ao princípio da livre apreciação da prova"[163], pecando, porém, na parte em que não se concedeu ao tribunal o acesso à informação sobre o funcionamento do algoritmo para a concretização plena de uma apreciação do resultado segundo as regras da experiência e a livre convicção.

[161] STARR, (2013), pp. 9-10. Trad. livre
[162] Num elenco não exaustivo, o art. 124.º CPP define o que pode constituir objeto da prova.
[163] LOBO, (2021), pp. 56-57.

Ainda que o princípio da livre apreciação da prova assuma um especial relevo na fase de julgamento[164], este princípio também vale para outras entidades judiciárias no decurso de todo o processo, nomeadamente para o juiz de instrução e para o MP[165] (também eles poderiam "dispor" dos juízos preditivos nos vários momentos de ponderação prognóstica, como vimos).

Reforça-se de novo, agora em senda probatória, a importância do princípio do contraditório, conforme previsto no art. 327.º/2 CPP, ao qual devem obedecer todas as provas, mesmo as que sejam produzidas oficiosamente pelo tribunal. Trata-se de uma disposição à qual será impossível de obedecer, se não houver "acesso ao código do algoritmo, ao seu claro funcionamento e à valoração de cada variável"[166]. Nesse sentido, "o controlo da codificação é um aspeto fundamental, já que a possibilidade de contestar os resultados fornecidos por um algoritmo é uma condição básica de um processo de decisão" que se quer *fair*.[167] Ainda assim, mesmo que o arguido tenha pleno acesso aos dados técnicos e à descrição do processo que levou àquele resultado, não é linear que a configuração desta dinâmica não venha criar outros inconvenientes no exercício do contraditório e dificultar o exercício do direito de defesa no geral. Pense-se, por exemplo, no encargo acrescido de ter de contestar exaustivamente todos os parâmetros do programa e o ónus desnecessário que recai sobre o arguido de provar, imagine-se, a título de exemplo, que o seu grau de escolaridade não é relevante, muito menos determinante, para considerá-lo uma pessoa mais ou menos perigosa.

Ainda neste campo, Hugo Luz dos Santos ressalva outra preocupação a respeito dos resultados "científicos" provenientes dos algoritmos preditivos: a vinculação do juiz[168]. A liberdade na apreciação da prova não significa discricionariedade na apreciação da prova. Esta é uma liberdade de acordo com o dever: dever de perseguir a realização da justiça e a descoberta da verdade material, pelo que tem de ser objetivável, motivável e suscetível de controlo[169]. Porém, caso o relatório de risco, dado a sua complexidade e tecnicidade

[164] Em especial porque, face ao princípio da imediação que supõe um contato imediato com os meios de prova, não valem para a formação da convicção do julgador quaisquer provas que não forem produzidas ou examinadas em audiência (art. 355.º/1 CPP). Cf. ANTUNES, (2022), pp. 211 e ss.
[165] ANTUNES, (2022), p. 202.
[166] CARIA, (2020), p. 255.
[167] RODRIGUES, (2020a), p. 22.
[168] SANTOS, (2022).
[169] ANTUNES, (2022), p. 202.

algorítmica[170], exija ser interpretado e explicado em tribunal por um perito,[171] estaremos perante uma prova pericial (art. 163.º CPP) cujo juízo técnico presume-se subtraído à livre convicção do julgador. No entanto, ainda que o juiz esteja vinculado ao parecer pericial na sua dimensão técnica, a avaliação jurídica da base de facto continua a pertencer-lhe em exclusivo[172], pelo que a proposta de "decisão" emitida pelo algoritmo inteligente terá de passar sempre pelo crivo do juiz titular do processo. De qualquer das formas, a última palavra pertence ao juiz, uma vez que só o ser humano, enquanto agente moral, pode ser responsabilizado pelas decisões tomadas[173].

A responsabilidade pela decisão relaciona-se estreitamente com a fundamentação da sentença (arts. 205.º/1 CRP e 374.º/2 CPP) que é exigida para se impor à generalidade das pessoas, e ao seu destinatário, o arguido, que tem de compreender o sentido das várias decisões que o afetam, sobretudo quando se trata de uma decisão de condenação, de forma a permitir o seu controlo. Isto obriga o juiz a desvendar a linha de raciocínio traçada para alcançar aquela decisão. Até porque, como afirma Anabela Rodrigues, "só uma decisão jurisdicional vinculada à lei e racionalmente justificada é essencialmente legitimadora do poder judicial"[174].

No entanto, tendo em conta a tendência humana para confiar nos resultados de um procedimento automatizado, "a submissão do juízo à máquina e a devolução da decisão sobre a punição ao algoritmo"[175] pode originar um efeito prejudicial contrário, chamado de "*automation bias*", manifestado quando as pessoas confiam em afirmações automaticamente gerada por computadores, mesmo quando confrontadas com a sua falta de rigor[176]. Esta prevalência quase inconsciente do resultado algorítmico sobre o juízo humano "arrisca-se a distrair o juiz da sua própria experiência profissional e a impedi-lo de fazer valer a discricionariedade na decisão"[177].

---

[170] Como acontece com os modelos de "caixa preta" que não são interpretáveis *by design* e requerem explicabilidade *post-hoc* para serem interpretados, através de técnicas conhecidas como "model induction". Cf. HUSSAIN, *et. al.*, (2021), p. 3.
[171] Segundo os arts. 151.º e ss. do CPP, a prova pericial tem lugar quando a perceção dos factos exige especiais conhecimentos técnicos, científicos ou artísticos.
[172] ANTUNES, (2022), p. 204.
[173] RODRIGUES, (2020a), pp. 43 e 52.
[174] RODRIGUES, (2020b), p. 272.
[175] RODRIGUES, (2020a), p. 45.
[176] CARIA, (2020), p. 259; *Ibid.*
[177] RODRIGUES, (2020a), p. 46.

V. A PERSPETIVA EUROPEIA – A CAMINHO DA REGULAÇÃO

1. A PROPOSTA DE REGULAMENTO DA INTELIGÊNCIA ARTIFICIAL

As críticas que comumente se tem apontado às tendências algorítmicas, e que supra sintetizamos, ganham nova relevância com a chegada da Proposta de Regulamento do Parlamento Europeu e do Conselho em matéria de IA[178] que veio reacender o debate sobre a contenção dos riscos desta tecnologia. Apresentada pela primeira vez a 21 de abril de 2021 e ajustada ao longo dos últimos meses, de modo a incluir o avanço das novas ferramentas de IA generativa, a proposta concretiza aquilo que, concluídas as negociações finais, poderá vir a ser a primeira Lei da Inteligência Artificial *("Artificial Intelligence Act"*) a nível comunitário, bem como um marco pioneiro a nível mundial. Na sua exposição de motivos, o diploma aponta como principal objetivo o desenvolvimento de um quadro jurídico único para uma IA segura, ética e de confiança por forma a garantir os direitos fundamentais e proteger os consumidores, sem prescindir do incentivo à inovação tecnológica.

Não obstante tratar-se de um documento extenso que procura acautelar diversas aplicações dos sistemas de IA, seja no que concerne a regras e requisitos sobre a sua colocação no mercado e utilização no espaço europeu, seja sobre as proibições de certas práticas, aqui iremos limitar-nos a fazer uma abordagem abreviada dos pontos gerais e destacar os preceitos que consideramos determinantes convocar para a discussão alusiva ao tema sob análise, precisamente, a admissibilidade de ferramentas preditivas no policiamento preditivo e nos processos de decisão judicial.

Antes de mais, o âmbito substantivo de aplicação do regulamento dedica-se aos *"sistemas de IA*", terminologia adotada para se referir a "um sistema baseado em máquinas concebido para operar com vários níveis de autonomia e que pode, em relação a objetivos explícitos ou implícitos, criar resultados, como previsões, recomendações ou decisões, que influenciam ambientes físicos ou virtuais" (art. 3.º/1 e consid. 6 da Proposta)[179].

[178] COM (2021) 206 final. Bruxelas, 21.4.2021 – 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.

[179] Todas as menções legislativas indicadas neste capítulo, salvo indicação diversa, referem-se a preceitos contidos na "Proposta de Regulamento (…) de IA", *Op. Cit.*

A proposta adotou uma abordagem regulatória em função do grau de risco que os sistemas de IA acarretam[180], tendo em conta o seu impacto na segurança e direitos fundamentais dos visados, ao categorizar os sistemas e aplicações de IA consoante quatro níveis de risco: inaceitável, elevado, limitado e mínimo.

A. RISCO INACEITÁVEL

A (proposta de) Lei[181] determina, no art. 5.º do Título II, que são proibidos, por representarem um risco inaceitável, os sistemas de IA que procuram distorcer ou manipular a consciência ou o comportamento humano[182]; os sistemas utilizados para fins de categorização biométrica e identificação biométrica à distância em "tempo real" ou "em diferido"[183] (*e.g.*, reconhecimento facial) em espaços públicos; "sistemas de IA para efeitos pontuação, avaliação ou classificação social de pessoas singulares ou grupos" com base no seu comportamento ou características pessoais, quando isso conduza a um tratamento prejudicial ou desfavorável das pessoas fora dos contextos nos quais os dados foram originalmente recolhidos; sistemas de IA concebidos para fazerem avaliações de risco individual ou de grupo, "a fim de avaliar o risco de uma pessoa singular poder cometer ou voltar a cometer uma infração ou para prever a ocorrência ou a recorrência de uma infração penal ou administrativa, real ou potencial", com base na avaliação de características pessoais, incluindo o perfil da pessoa, a sua localização ou comportamento criminoso anterior, bem como sistemas de IA para fins de reconhecimento de emoções ou movimentos fisiológicos na aplicação do direito, controlo de fronteiras, locais de trabalho e instituições de ensino.

Observa-se que a seleção de práticas perigosas ali compiladas relaciona-se com usos particularmente nocivos dos sistemas de IA que são suscetíveis de conflituarem com os valores fundamentais da União – como o respeito pela dignidade humana, liberdade,

[180] Sem prejuízo das proibições aplicáveis quando uma prática de IA infringe outra legislação, incluindo o acervo da UE em matéria de proteção de dados, proteção do consumidor ou concorrência.
[181] Atende-se à análise do Regulamento na sua proposta inicial (COM (2021) 206 final) com as Alterações aprovadas pelo Parlamento Europeu, em 14/06/2023, 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 (P9_TA (2023) 0236).
[182] Exceto quando sejam destinados a fins terapêuticos aprovados com base no consentimento informado específico do indivíduos a eles expostos.
[183] Consiste na análise de imagens gravadas e só será admissível para a persecução de infrações penais graves que já tenham ocorrido e mediante autorização judicial. Cf. consid. 8 e art 5.º/1/e).

igualdade, democracia e Estado de direito, direito à não discriminação, proteção de dados e privacidade – pelo que se proíbe de antemão a sua introdução e utilização na Europa. É nesta categoria que se incluem as aplicações de sistemas de IA para fins de policiamento preditivo na vertente das técnicas orientadas para a identificação individual, i.e., quando procuram prever potenciais futuros autores de crimes.

Esta proibição é corroborada pelo RGPD, onde, atendendo às preocupações quanto à natureza dos dados utilizados e a respetiva interferência na privacidade dos sujeitos, o artigo 22.º/1 salvaguarda o direito do titular dos dados a não *ser "sujeito a nenhuma decisão tomada exclusivamente com base no tratamento automatizado, incluindo a definição de perfis, que produza efeitos na sua esfera jurídica ou que o afete significativamente de forma similar".* Descartando-se desta forma, a possibilidade de se criarem perfis algorítmicos, numa espécie de "*scoring* de risco" ou "*scoring* de perigosidade" dos sujeitos à disposição das autoridades policiais.

O art. 5.º/1/d-A) cinge o seu âmbito a pessoas singulares ou grupos, deixando, porém, de fora os sistemas de IA vocacionados para as áreas de policiamento como é o caso das práticas orientados para identificar os "*hot spots*" da ocorrência de crimes que não necessitam necessariamente de utilizar informações de identificação pessoal[184] – sem embargo desses sistemas poderem não ser totalmente inócuos dado a sua tendência para sustentar discriminações sistémicas, como anteriormente apontado. Pois, ainda que se mencione a proibição para "prever a ocorrência ou a recorrência de uma infração penal", esta é sempre por referência à utilização de perfis individuais. Mesmo assim, este preceito também não parece incluir, ainda que dentro da aplicação de sistemas de IA direcionadas para a identificação individual, aqueles que visem identificar potenciais vítimas de infrações penais, hipótese que irá, no entanto, recair nas práticas de risco elevado (Anexo III, ponto 6, al. f)).

[184] Ferramentas destinadas a prever crimes são geralmente alimentadas com dados histórico, em grande parte de fontes oficiais, sobre o local e o tipo de crimes cometidos, incluindo variáveis sobre o contexto como a densidade populacional ou a ocorrência de grandes eventos. Cf. European Union Agency for Fundamental Rights (FRA), (2020). "Getting the future right. Artificial intelligence and fundamental rights", pp. 34-35.

B. RISCO ELEVADO

O art. 6.º do título III, estabelece as regras de classificação dos sistemas de IA de "risco elevado" e identifica duas categorias principais: (1) os sistemas de IA concebidos para serem utilizados como componentes de segurança de um produto ou serem eles próprios produtos abrangido pela legislação enumerada no anexo II e sujeitos a uma avaliação da conformidade *ex ante* por terceiros; (2) outros sistemas de IA abrangidos por um ou mais domínios críticos e casos de utilização[185] referidas no anexo III que, à luz da finalidade a que se destinam, representam, ou são suscetíveis de representar, um dano significativo para a saúde, a segurança e os direitos fundamentais das pessoas físicas na União e, caso sejam aplicados numa infraestrutura crítica, também um dano para o meio ambiente. Tal risco significativo de dano deve ser avaliado em relação ao efeito que o risco tem em função do "seu nível de gravidade, intensidade, probabilidade de ocorrência e duração" e o facto de poder afetar um indivíduo ou uma pluralidade de pessoas (consid. 32). Portanto, um sistema de IA é identificado como de alto risco e, por isso, sujeito aos encargos regulatórios adicionais, apenas quando, em função das suas características ou campo em que se aplicação, acarreta um impacto prejudicial substancial para os interesses protegidos.

No entanto, trata-se de uma presunção ilidível que admite prova em contrário, uma vez que o próprio regulamento permite que os fornecedores, cujos sistemas de IA se enquadrem nos usos listados no Anexo III, que considerem que o seu sistema não representa um risco significativo para os interesses mencionados, e que, portanto, não devem ser sujeitos aos requisitos legais aplicáveis a esta categoria de risco, devem informar, antes da colocação do sistema no mercado ou da sua entrada em serviço, a autoridade nacional de controlo[186] (ou o Gabinete Europeu de Inteligência Artificial – *"AI Office"* – caso o sistema de IA se destine a ser utilizado em dois ou mais Estados-Membros) mediante a apresentação de uma declaração fundamentada que descreva as informações relevantes do sistema em questão e as razões pelas quais este não representa um risco significativo (consid. 32-A e art. 6.º/2-A). Abre-se aqui uma fresta importante que permite aos fornecedores dos sistemas de

[185] A saber: identificação biométrica e categorização de pessoas singulares; gestão e funcionamento de infraestruturas críticas; 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; manutenção da ordem pública; gestão da migração, do asilo e do controlo das fronteiras; administração da justiça e processos democráticos.

[186] Autoridade pública nacional responsável pela execução e aplicação do regulamento em cada Estado-Membro (art. 3.º/42).

IA que atuem nos domínios destacados desmarcarem-se previamente do cumprimento dos requisitos prescritos para os sistemas de risco elevado. Caberá à Comissão especificar os critérios que irão permitir às empresas avaliar se o seu sistema representaria tais riscos, bem como desenvolver um modelo único para a notificação (consid. 32-A).

O Anexo III é especialmente importante porque os sistemas de IA aqui mencionados são considerados "casos de utilização críticos" e estarão sujeitos ao escrutínio máximo imposto pela Lei. Para o escopo deste trabalho, interessam-nos particularmente os números 6 e 8 do Anexo III. O número 6 refere-se aos sistemas de IA usados pelas autoridades policiais na manutenção da ordem pública, isto é, para efeitos de prevenção, deteção, investigação e repressão de infrações penais. Quando aplicados neste âmbito, a lei considera serem de risco elevado, desde que preencham os critérios do art. 6.º/2, sistemas como polígrafos e instrumentos semelhantes (al. b)); sistemas utilizados para avaliar a fiabilidade das provas (al. d)); para a definição de perfis[187] de pessoas singulares (al. f)) e para pesquisar grandes conjuntos de dados complexos, disponíveis em diferentes fontes e formatos de dados com o intuito de identificar padrões desconhecidos ou descobrir conexões ocultas nos dados (al. g)).

Como mencionado supra, o exemplo de aplicação que encaixaríamos nesta última alínea, precisamente pela extensa análise de dados com recurso a sistemas de IA que pode envolver, seria, desde logo, o policiamento das localidades identificadas como palco dos próximos crimes mas, também, num sentido mais preventivo e repressivo do crime, terão aqui cabimento os sistemas aplicados em domínios criminais especializados que procuram por indícios de crimes, dos seus autores ou vítimas, dentro de uma base delimitada de atuação designadamente nos contextos do cibercrime, pornografia infantil ou crimes económicos. A título ilustrativo apontamos o caso do sistema "Connect"[188] – principal ferramenta usada pela autoridade tributária e aduaneira do Reino Unido na análise de transações financeiras e deteção de irregularidades tendo em vista o combate à fraude e evasão fiscal – e, ainda no

[187] Na aceção do artigo 3.º/4, da Diretiva (UE) 2016/680. A saber, "qualquer forma de tratamento automatizado de dados pessoais que consista em utilizar esses dados pessoais para avaliar certos aspetos pessoais de uma pessoa singular, nomeadamente para analisar ou prever aspetos relacionados com (…) interesses, fiabilidade, comportamento, localização ou deslocações".
[188] Mais em: https://www.taxation.co.uk/articles/hmrc-s-connect-computer-and-investigations

espaço digital, o exemplo da base de dados "International Child Sexual Exploitation" (ICSE)[189] gerida pela Interpol para localizar vítimas de abuso sexual infantil[190].

Já o número 8 do Anexo III apresenta os usos na administração da justiça e em processos democráticos que, "tendo em conta o seu impacto potencialmente significativo na democracia, no Estado de direito, nas liberdades individuais, bem como no direito à ação e a um tribunal imparcial", devem ser classificados como sendo de risco elevado (consid. 40). Aqui incluem-se sistemas de IA concebidos para serem usados pelas autoridades judiciárias ou órgãos administrativos, ou em seu nome, no auxílio à investigação e interpretação dos factos e do direito, bem como na aplicação da lei a um conjunto específico de factos ou usada de maneira semelhante na resolução alternativa de litígios.

Observa-se, que o regulamento em si, bem como o número 8 do Anexo III, referem-se a ambientes judiciais no geral sem particularizar os usos ou consequências específicas que a IA pode suscitar ao nível da justiça criminal, nem tão pouco discriminam o seu emprego nas diferentes fases processuais.

Outros usos de risco elevado previstos remetem a sistemas de IA destinados a influenciar o resultado ou o exercício de voto das pessoas físicas de uma eleição ou referendo, exceto as ferramentas para organizar e otimizar campanhas políticas do ponto de vista administrativo e logístico; e sistemas de IA destinados a serem usados por plataformas de redes sociais nos seus sistemas de recomendação.

Contudo, tal qualificação "não deve ser alargada aos sistemas de IA destinados a atividades administrativas puramente auxiliares que não afetem a administração real da justiça em casos individuais, como a anonimização ou pseudonimização de decisões judiciais, documentos ou dados, comunicação entre pessoal, tarefas administrativas ou afetação de recursos" (consid. 40). Estas aplicações de suporte à atividade jurisdicional são as que atualmente predominam no uso judiciário, referidas em vários exemplos ao longo deste trabalho. E ainda que não cobertas pelas exigências aplicáveis às ferramentas de alto risco, serão, certamente, sujeitas aos princípios gerais do art. 4.º-A – extensível a todos os sistemas de IA independentemente do risco, no sentido de se promover "uma abordagem europeia centrada no ser humano coerente para uma inteligência artificial ética e fiável", a

[189] https://www.interpol.int/Crimes/Crimes-against-children/International-Child-Sexual-Exploitation-database
[190] Cf. CEPEJ (2018) 14. "Carta Europeia de Ética (…)", *Op. cit.*

ter em conta: controlo e supervisão por humanos, robustez técnica e segurança, privacidade e governança de dados, transparência, diversidade, não discriminação e equidade, bem-estar social e ambiental.

Mais adiante, em resultado da hierarquização do risco, o legislador definiu, nos arts. 8.º a 15.º do Capítulo 2, Título III, os requisitos específicos obrigatórios para a admissão de sistemas de IA de risco elevado, nomeadamente, sistema de gestão de riscos, exigências sobre a documentação técnica, o registo de eventos do sistema, exigências sobre a transparência e prestação de informações aos utilizadores. Já o Capítulo 3 concretiza as obrigações dos fornecedores e utilizadores de sistemas IA de alto risco e outras partes envolvidas, das quais salientamos a elaboração de uma declaração de conformidade UE (arts. 16.º/e-A) e 48.º e Anexo V) e o registo obrigatório do sistema de IA na base de dados da UE (arts. 16.º/f), 51.º e 60.º). Estes requisitos devem assegurar que os sistemas de IA não representam riscos inaceitáveis para interesses públicos importantes da União (consid. 27).

Desta forma, a responsabilidade pelo cumprimento das condições de admissibilidade e funcionamento de uma ferramenta de IA de risco elevado é assumida pelo fornecedor num procedimento de autoavaliação *ex-ante*. Depois de efetuar a avaliação da conformidade necessária e elaborar a respetiva declaração, antes da colocação do sistema no mercado (consid. 62 e art. 48.º), o fornecedor deve registar esses sistemas de IA autónomos na base de dados da UE (consid. 69) que será gerida pela Comissão, a fim de aumentar a transparência e permitir a supervisão *ex post* por parte das autoridades competentes[191].

Aqui surge uma crítica pertinente relacionada com a opção política de se optar por uma abordagem de "autoexecução" das exigências da conformidade *ex-ante* que pode suscitar preocupações com a falta de regulamentação, uma vez que a estrutura de (auto) avaliação proposta depende principalmente de controlos internos conduzidos pelo fornecedor de IA.[192] Isto porque, a exigência de uma mera avaliação interna para os casos do art. 43.º/2, pensada com um fornecedor médio diligente em mente, pode vir a ser preocupante se considerada sem mais. Na maioria dos casos, a lei não requer supervisão

[191] Ponto 5.2.3. da Exposição de motivos da Proposta.
[192] EBERS, Martin, *et. al.*, (2021). "The European Commission's Proposal for an Artificial Intelligence Act – A Critical Assessment by Members of the Robotics and AI Law Society (RAILS)". 4(4), pp. 589-603. https://doi.org/10.3390/j4040043; RAPOSO, Vera L., (2022). "Ex machina: preliminary critical assessment of the European Draft Act on artificial intelligence", *International Journal of Law and Information Technology*, Vol. 30, Issue 1, p. 98. https://doi.org/10.1093/ijlit/eaac007

externa prévia, deixando a avaliação à responsabilidade do fornecedor. O envolvimento de um organismo terceiro, num controle *ex ante* independente, é convocado apenas para um número limitado de situações (cf. considerando 64).

C. RISCO LIMITADO E RISCO MÍNIMO

Já a categoria de sistemas de IA de "risco limitado" surge para incluir sistemas específicos definidos como "sistemas de IA de uso geral" [193] e "modelos de base"[194] onde se podem encaixar sistemas de produção de conteúdo "*deepfake*" e os programas de conversação automática de IA generativa que abrangem os modelos de linguagem em grande dimensão (*e.g.,* ChatGPT e Bard). Estes sistemas ficarão sujeitos a requisitos de conformidade mais leves o que significa que, para além dos princípios gerais, devem cumprir os requisitos mínimos de transparência (art. 28.º-B/4/a)), nomeadamente publicar informações sobre o uso de dados de treino que possam estar protegidos por direitos de autor. Ademais, as pessoas devem ser informadas de que estão a interagir com um sistema de IA e que o conteúdo está a ser manipulado artificialmente ou gerado por meios automatizados e não por humanos, sob reserva de exceções para fins legítimos[195], de modo a permitir uma tomada de decisão mais informada[196].

Por último, no risco mínimo cairão os sistemas de IA que não se enquadrem nas três categorias mencionadas e, por isso, não estarão sujeitos a cumprir quaisquer obrigações legais adicionais, mas podem vir a ter um código de conduta próprio que fomente a sua conformidade com os demais requisitos, de forma voluntária (art. 69.º).

---

[193] Trata-se de sistemas que podem executar diferentes funções como o reconhecimento de imagem e fala, geração de áudio e vídeo, deteção de padrões, resposta a perguntas, tradução e outros, e podem ser aplicáveis em diversos contextos ou integrados noutros sistemas para os quais não foram inicialmente projetados. Consid. 70 e art. 3.º/1-D e 44-D.

[194] São treinados com dados em escala e podem ser adaptados a uma ampla gama de tarefas (consid. 60-E).

[195] Consubstancia-se como exceções a manutenção da ordem pública e o exercício do direito à liberdade de expressão e liberdade das artes, ou seja, não se aplica a sistemas de IA para fins recreativos e quando estejam legalmente autorizados a detetar, prevenir, investigar e reprimir infrações penais, salvo se esses sistemas estiverem disponíveis ao público para denunciar uma infração penal. Cf. art. 52.º/3.

[196] Exposição de motivos, ponto 5.2.4 e arts. 28.º-B/4 e 52.º/1.

D. CONSIDERAÇÕES FINAIS

Da análise entendemos que o critério para classificar um sistema como de risco elevado atende não só à natureza potencialmente perigosa do sistema de IA, mas depende também das modalidades e da finalidade específica para as quais aquele sistema é utilizado. Resulta da Proposta que estes sistemas são admitidos com a condição de cumprimento dos requisitos obrigatórios constantes do capítulo 2, Título III. Esses requisitos devem assegurar que os sistemas de IA disponíveis na União não representam riscos inaceitáveis para importantes interesses públicos, tal como reconhecidos e protegidos pelo direito da União, incluindo os direitos fundamentais, a democracia, o Estado de direito ou o meio ambiente. Desta forma, estes só serão considerados de alto risco se comprometerem os direitos supramencionados, exceto se o fornecedor declarar a "inofensividade" do seu sistema antes de colocá-lo à disposição no espaço europeu. Porém, não é certo que numa avaliação prévia seja possível detetar cabalmente todos os potenciais riscos que possam surgir em contextos reais de utilização ou a longo prazo.

Analisando os preceitos com os algoritmos preditivos em mente, e para o que aqui importa, depreendemos que as avaliações de risco individual sobre a possibilidade de o agente cometer uma infração ou reincidir criminalmente, no sentido que estas têm vindo a ser empregues pelas forças policiais e tribunais estadunidenses, serão agora proibidas no geral, dentro do espaço europeu, independentemente do campo em que se apliquem (art. 5.º), colocando, assim, um entrave às aplicações preditivas de risco que têm o suspeito ou arguido como destinatário, Dizemos "agora" porque não se pode deixar de mencionar a intenção diversa que vinha contida na primeira versão da Proposta[197] no sentido da admissão, na categoria de risco elevado, dos sistemas de IA utilizados em avaliações de risco individuais para prevenir, investigar, detetar ou processar uma infração penal, posição que não logrou estabelecer-se, muito por força da pressão de organizações defensoras dos direitos civis[198].

[197] Cf. ponto 6 do Anexo III, em COM (2021) 206 final de 21.04.2021.

[198] Entre outros apelos, a Fair Trials, European Digital Rights (EDRi) e 49 outras organizações lançaram uma declaração coletiva para pedir à UE que proíba os sistemas de policiamento preditivo com base nos argumentos da lesão de direitos fundamentais face à discriminação, vigilância e excesso de policiamento; violação do direito à liberdade, ao direito a um julgamento justo e à presunção de inocência, e a falta de transparência, responsabilidade e direito a um recurso efetivo. Disponível em: https://www.fairtrials.org/articles/news/ai-act-eu-must-ban-predictive-ai-systems-in-policing-and-criminal-justice/

Na versão atual, ainda que a lei tenha desejado restringir a classificação de risco elevado a certos âmbitos de aplicabilidade, a delimitação prática da fronteira que circunscreve os sistemas que aqui se enquadram continua muito ampla, sobretudo, no contexto da aplicação da lei. Vejamos, então, algumas funções que poderão vir a ser desempenhadas pela IA à luz desta regulamentação. No cenário judicial, os sistemas de IA serão admitidos, sob apertados requisitos da categoria de risco elevado, quando sirvam fins de auxílio à investigação e interpretação dos factos e do direito e de aplicação da lei a um conjunto específico de factos.

Em referência à investigação criminal, convocamos de imediato os programas forenses de IA usados durante o inquérito enquanto meio de obtenção de prova em ambiente digital, que normalmente se associa a métodos ocultos de investigação, designadamente através do *malware*, ou mesmo enquanto método aberto (*e.g.*, durante uma pesquisa informática).[199] Ainda na senda probatória, a IA quando aplicada à prova pericial, que muitas vezes constitui uma causa de demora no andamento do processo, pode ser útil num apuramento mais ágil de "questões técnicas de cujo conhecimento dependa a descoberta da verdade material"[200]. Pense-se, em sistemas de IA que se destinam a verificar a fiabilidade de um meio de prova, por exemplo, a autenticidade de um documento apresentado no processo[201], ou a produzir uma reconstituição exata do facto, *e.g.*, "as condições em que ocorreu um acidente de viação"[202]. No que concerne à assistência na interpretação jurídica e na aplicação da lei a um conjunto específico de factos, parece-nos ser este o preceito que dá alguma abertura (ainda que coberta pelos requisitos da categoria de alto risco) para introduzir sistemas de IA nos processos de tomada de decisão, no geral, e na fase determinante do processo judicial, em especial. Desde logo, para selecionar e identificar os factos processuais relevantes[203] para o controlo dos pressupostos processuais e subsequente verificação de eventuais exceções dilatórias – *e.g.*, "identificação automática de situações em que se verifica a incompetência absoluta do tribunal"[204].

---

[199] Desenvolvido em FIDALGO, Sónia (2020). "A utilização de inteligência artificial no âmbito da prova digital – direito fundamentais (ainda mais) em perigo", pp. 129-161, in RODRIGUES, Anabela M. (Coord.), (2020). «A Inteligência Artificial no Direito Penal», Coimbra: Almedina.
[200] GONÇALVES, (2022), p. 277.
[201] A produção de prova documental em processo penal, incluindo a respetiva falsidade, rege-se pelos arts. 164.º a 170.º CPP.
[202] Exemplo de GONÇALVES, (2022), p. 277.
[203] Na fase de julgamento, seriam os factos objeto do processo descritos na acusação.
[204] GONÇALVES, (2022), p. 278.

Também no julgamento, à parte da utilização de ferramentas de risco para determinar a perigosidade ou risco de reincidência de um arguido (que a final se proíbe), outras funções podem ser atribuídas aos sistemas de IA com consequências na determinação da pena (em sentido amplo) numa dinâmica de cooperação entre o juiz e o sistema algorítmico. Abandonando-se a lógica preditiva e agora com o foco no desempenho de uma função mais analítica, pense-se na possibilidade de incutir nos sistemas de IA parâmetros mensuráveis e matematizáveis sobre os factos dados como provados, designadamente os que dizem respeito à punibilidade (*e.g.*, verificação das condições objetivas de punibilidade) e à existência de circunstâncias modificativas, de forma a ser possível deduzir a moldura da pena aplicável[205] de um modo automático. Ou a possibilidade de, descritos os factos, a IA auxiliar na determinação de outras questões onde "a decisão assente em critérios objetivos ou em operações de simples cálculo aritmético"[206] possa ser processada de forma mais eficaz por sistemas algorítmicos. Aqui pode entrar a operação de determinação de um quantitativo diário adequado da pena de multa (art. 47.º/2 CP) que deve ser encontrada atendendo à "situação financeira do condenado" cujos critérios de apuração a lei não define[207], mas a capacidade analítica dos sistemas de IA pode prontamente calcular.

A propósito da assistência que estes sistemas podem prestar ao juiz nos vários momentos, o considerando 40 reforça a ideia intangível de que "a utilização de ferramentas de inteligência artificial pode auxiliar mas não deve substituir o poder de decisão dos magistrados ou a independência judicial, uma vez que a decisão final deve continuar a ser uma atividade humana."[208] Ademais, é notório um esforço particular da proposta em garantir que o funcionamento destes sistemas seja suficientemente transparente para os utilizadores.

Assegura-se ainda que os responsáveis pela implantação de sistemas referidos no anexo III que tomam ou ajudam a tomar decisões relacionadas com pessoas singulares devem informá-las de que estão a ser sujeitas à utilização de um sistema de IA de risco elevado (art. 29.º/6-A). Este direito a uma explicação sobre a decisão individual deve refletir uma explicação clara sobre a finalidade, os principais parâmetros da decisão e os dados que

[205] ANTUNES, Maria J. (2023). «Penas e Medidas de Segurança», Coimbra: Almedina, 2ª ed, p. 49-52.
[206] GONÇALVES, (2022), p. 279.
[207] ANTUNES, (2023), p. 59.
[208] Consid. 40; Também em (2020/2016(INI)), "A inteligência artificial no direito penal (…)", *Op. Cit.*, p. 9, onde o Parlamento Europeu já apelava à proibição do uso de IA e de tecnologias relacionadas para propor decisões judiciais, de modo a preservar "o poder discricionário soberano dos juízes e a tomada de decisões numa base casuística" – ponto 16.

foram usados (art. 68.º-C), podendo ainda o titular dos dados que esteja a ser alvo de uma decisão totalmente automatizada que o afete diretamente chamar à colação o disposto no n.º 3 do art. 22.º do RGPD e "obter intervenção humana por parte do responsável, manifestar o seu ponto de vista e contestar a decisão".

No que toca às consequências previstas, as sanções aplicáveis em caso de infração ao presente regulamento devem ser estabelecidas pelos Estados-Membros, podendo as coimas, em casos de incumprimento mais grave, i.e., violação de uma prática proibida, alcançarem valores até os 40 milhões de euros ou, se o infrator for uma empresa, até 7% do volume de negócios (art. 71.º e ss.). Neste sentido, a proposta ao pretender essencialmente a responsabilização dos produtores, poderá ter, em contrapartida, um efeito dissuasor no que toca ao estabelecimento de pequenas e médias empresas no mercado europeu, com as devidas repercussões que isso acarreta para a competitividade da UE no plano da inovação e desenvolvimento tecnológico. Parece-nos natural que para cumprir as novas exigências, mesmo as empresas dominantes na indústria da tecnologia (*Big Tech*), terão de suportar encargos acrescidos para se adaptarem e adequarem os seus processos de criação, produção e comercialização à nova regulamentação. O impacto financeiro, a sujeição a exigências de conformidade suplementares, bem como as pesadas multas administrativas em caso de incumprimento, são fatores que podem pesar na decisão aquando da ponderação das empresas em investirem no mercado europeu.

Por fim, ainda que se espere uma versão definitiva do Regulamento até ao final deste ano, a Comissão Europeia admite ser expectável que as regras só entrem em vigor em 2025. Pelo que, não obstante prever-se uma atualização constante dos preceitos[209], são vários os receios que antecipam que este se possa tornar, aquando da sua entrada em vigor, num projeto já ultrapassado pela sobrepujante velocidade da inovação. Nesta senda, Luís Paulo Reis, presidente da Associação Portuguesa para a IA, critica a tempestividade e adequação deste regulamento aos tempos hodiernos, considerando tratar-se de uma proposta "absolutamente desatualizada", sustentando que as medidas contempladas "já não têm nada a ver com os sistemas de IA que temos atualmente"[210].

[209] Atendendo ao acelerado desenvolvimento tecnológico e a imprevisibilidade de riscos futuros, os arts. 5.º, 52.º e a lista de domínios e usos de risco elevado do Anexo III devem estar sujeitos a revisão e atualização regular (consid. 27, 85-A e art. 7.º/1).

[210] FERREIRA, Filipa & CAPUCHO, Luísa (2023). "Inteligência Artificial: "A tecnologia é imparável", mas precisa de "regulamentação adequada"". *JPN*. https://www.jpn.up.pt/2023/06/16/inteligencia-artificial-a-tecnologia-e-imparavel-mas-precisa-de-regulamentacao-adequada/

VI. CONCLUSÕES

O acelerado desenvolvimento tecnológico verificado nas últimas décadas e a crescente expansão do campo da Inteligência Artificial, movida a *Big Data*, *machine learning* e outras técnicas computacionais, têm suscitado mudanças substanciais em praticamente todas as áreas da nossa vida. Os desafios no âmbito do direito penal não foram exceção. Aqui, para além das oportunidades que se abrem ao nível de uma nova criminalidade potencializada pelos mecanismos da IA – numa perceção da tecnologia como uma "faca de dois gumes" que pode também servir causas mal-intencionadas –, cria-se um impacto significativo na forma como se contempla e realiza a prevenção, deteção, investigação e repressão das infrações penais. Deixámos, no entanto, os fins criminosos de lado, para focar em aplicações consideradas legítimas que pudessem beneficiar das capacidades da IA na persecução dos seus objetivos, com destaque para a sua utilização na manutenção da ordem pública e no plano da administração da justiça.

Numa abordagem central, procurámos explorar as vantagens e os perigos que podem resultar da aplicação dos resultados gerados por sistemas ditos "inteligentes" em áreas particularmente melindrosas do ponto de vista do confronto com os direitos fundamentais das pessoas. No caso do policiamento preditivo, destacamos o perigo de resultados tendenciosos ou erróneos e as ameaças à privacidade em consequência do acesso e análise de dados pessoais em massa. Já no plano processual, os ganhos de objetividade e precisão que podem resultar dos cálculos algorítmicos nos momentos decisórios, acarretam também o risco de suprimir a necessária discricionariedade do juiz na apreciação das questões jurídico-penais e podem prejudicar o exercício do direito de defesa.

Através desta exposição procurámos olhar para a IA de uma perspetiva prognóstica, visto que um dos objetivos da justiça é tornar-se o mais previsível possível a fim de promover a segurança jurídica. Durante este trajeto, enunciamos uma panóplia de riscos que não pode ser desconsiderada num debate sobre a admissibilidade de algoritmos de IA, no geral, e na justiça em particular, que acreditamos também ter sido levada em consideração nas motivações que estiveram na origem da Proposta de Regulamento da UE para a IA. Trata-se de um diploma que procura mitigar os perigos desta tecnologia e que vai ditar o seu futuro próximo na Europa, numa disposição de contornos ainda não totalmente claros para a maioria dos operadores – isto porque, para além da aprovação no processo legislativo final,

ainda está dependente da concretização de várias etapas práticas (*e.g.*, criação dos organismos e estruturas de controlo, implementação da base de dados de registo único e da densificação de certos parâmetros deixados ao critério da Comissão), que se preveem no diploma e vão, de forma articulada, executar o plano traçado para cumprir com a ambição de tornar a IA responsável, confiável, segura e ética para os utilizadores da União.

Num momento de ponderação legislativa, parece-nos que o regulamento optou essencialmente por um ponto de equilíbrio entre a contenção de certos usos ou finalidades de aplicação, considerando os efeitos que podem surgir sobre direitos fundamentais, e um quadro regulatório que fomente um desenvolvimento tecnológico responsável e, tanto quanto possível, seguro. Sendo explícito que não se trata de uma tecnologia totalmente inócua, é particularmente de saudar uma abordagem de cautela vertida na Proposta e focada numa qualificação gradual do risco dos sistemas de IA num cruzamento tripartido entre alguns fins específicos de aplicação, certas práticas e o risco de dano que delas pode resultar.

Dos pontos essenciais retirados deste diploma, destacamos que a utilização de instrumentos preditivos de IA que visem calcular o risco que um sujeito demonstra para cometer uma infração penal ou reincidir criminalmente, será totalmente proibida, em qualquer âmbito de aplicação – inclusive no policiamento preditivo e na tomada de decisão judicial. A lei, no entanto, reconhece a utilidade e admite a aplicabilidade dos sistemas de IA como ferramentas de auxílio na administração da justiça, desde que isso não implique a substituição total do decisor humano. Portanto, sistemas que procurem assessorar nas tomadas de decisão de quem investiga, acusa e julga, não serão, à partida, proibidos, mas a sua admissibilidade dependerá do cumprimento pelos fornecedores desta tecnologia de certas condições, previstas para os sistemas da categoria de risco elevado.

Desta forma, a UE, com os valores da transparência e segurança em mente, apenas estabelece os limites intransponíveis que se devem verificar no processo de implementação e utilização das tecnologias de IA no espaço europeu, mas caberá a cada Estado-Membro, atendendo às especificidades do ordenamento jurídico, fazer as suas opções políticas no sentido de estabelecer até onde será benéfica essa introdução.

No caso de Portugal, os sistemas de IA ainda têm pouca expressão prática, mas são notórios os esforços perpetuados a caminho de uma maior digitalização dos serviços públicos e jurídicos, com o objetivo de diminuir a morosidade dos atos, facilitar o acesso aos materiais, garantir a segurança processual dos documentos e dinamizar o trabalho dos

operadores judiciais. Dadas as suas vantagens, a crescente incorporação de recursos de IA no campo forense é um movimento que nos parece inevitável, tendo já registado comprovados ganhos de celeridade, eficácia e acessibilidade. No entanto, a nosso ver, as próprias especificidades da atividade jurídica são o que constituem, neste momento, o maior impedimento para estender a aplicação da IA à tomada de decisão judicial, não obstante poder coadjuvar no "cálculo" de quesitos menores, acessórios à decisão humana principal, na medida que o próprio Regulamento o permitir.

Todavia, consideramos que o sucesso de uma implementação pacífica e responsável desta tecnologia em deliberações determinantes depende, não só da conformidade com as orientações da UE, mas, também, de uma aceitação generalizada por parte da sociedade em delegar o peso da decisão para uma máquina, no limite, prescindindo do discernimento do julgador humano, o que exige uma inabalável confiança na segurança e *fairness* dos algoritmos, que cremos ainda não ter sido alcançada.

"Machines have calculations. We have understanding.
Machines have instructions. We have purpose.
Machines have objectivity. We have passion.
(...)
And if we fail, it's not because our machines are too intelligent, or not intelligent enough. If we fail, it's because we grew complacent and limited our ambitions. Our humanity is not defined by any skill, like swinging a hammer or even playing chess. There's one thing only a human can do. That's dream. So let us dream big."[211]

[211] Garry Kasparov durante a conferência TED Talks: "Don't fear intelligent machines. Work with them" sobre a tecnologia que o "derrotou" há 25 anos. Disponível em: https://www.youtube.com/

VII. BIBLIOGRAFIA

ANTUNES, Maria J., (2022). «Direito Processual Penal», Coimbra: Almedina, 4ª ed.

ANTUNES, Maria J., (2023). «Penas e Medidas de Segurança», Coimbra: Almedina, 2ª ed.

BURCHARD, Christoph, (2021). "Artificial Intelligence as the End of Criminal Law? On the Algorithmic Transformation of Society", pp. 165-205, *in* ANTUNES, Maria J. & SOUSA, Susana A. de (Eds.), (2021). «Artificial Intelligence in The Economic Sector Prevention and Responsibility», Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra,

COPELAND, B. Jack, (2000). "The Turing Test. Minds and Machines", 10, pp. 519–539. https://doi.org/10.1023/A:1011285919106

EBERS, Martin.; HOCH, Veronica R. S.; ROSENKRANZ, Frank; RUSCHEMEIER, Hannah & STEINRÖTTER, Björn, (2021). "The European Commission's Proposal for an Artificial Intelligence Act – A Critical Assessment by Members of the Robotics and AI Law Society (RAILS)". 4(4), pp. 589-603. https://doi.org/10.3390/j4040043

ENGELMANN, Wilson & FRÖHLICH, Afonso V. K., (2020). "Inteligência Artificial Aplicada À Decisão Judicial: O Papel dos Algoritmos no Processo de Tomada de Decisão". *Revista Jurídica* (FURB), Vol. 24, N.º 54.

FERRARI Isabela; BECKER Daniel & WOLKA Erik N., (2018). "Arbitrium ex machina: panorama, riscos e a necessidade de regulação das decisões informadas por algoritmos". *Revista dos Tribunais*. Vol. 995.

GUIMARÃES, Rodrigo R. C., (2019). "A Inteligência Artificial e a disputa por diferentes caminhos em sua utilização preditiva no processo penal". *Revista Brasileira de Direito Processual Penal*, Porto Alegre, Vol. 5, N.º 3, p. 1555-1588. https://doi.org/10.22197/rbdpp.v5i3.260

HAENLEIN, Michael; KAPLAN Andreas, (2019). "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence", *Business Horizons*, Vol. 62 (1), pp. 15-25. https://doi.org/10.1016/j.bushor.2018.08.004

HUSSAIN, Fatima, HUSSAIN, Rasheed & HOSSAIN, Ekram, (2021). "Explainable Artificial Intelligence (XAI): An Engineering Perspective". https://doi.org/10.48550/arXiv.2101.03613

LOBO, Fernando R., (2021). "A utilização de sistemas preditivos de inteligência artificial na justiça", Lusíada, Revista de Direito 23/24, pp. 49-64.

MCCARTHY, John; MINSKY, Marvin L., ROCHESTER, Nathaniel & SHANNON, Claude E., (1955). "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence". http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

MCCARTHY, John, (2007). *«What is AI? / Basic Questions»*, University of Standford, pp. 1 – 15. http://jmc.stanford.edu/articles/whatisai.html

NOVAIS, Paulo & FREITAS, Pedro M., (2018). "Inteligência Artificial e Regulação de Algoritmos", *Diálogos, União Europeia-Brasil*.

OLIVEIRA, Arlindo, (2019). «Inteligência Artificial», Fundação Francisco Manuel dos Santos.

PEDRINA, Gustavo M. L., (2019). "Consequências e perspectivas da aplicação de inteligência artificial a casos penais", *Revista Brasileira de Direito Processual Penal*, Porto Alegre, set.-dez., Vol. 5, n.º. 3, pp. 1589-1606.

PEREIRA, Alexandre L. D., (2020). "Inteligência Artificial na Decisão Jurisprudencial?". *JURISMAT*, Portimão, n.º 12, pp. 73-92.

PEREIRA, Alexandre L. D., (2019). "O Responsável pelo Tratamento de Dados Segundo o RGDP". Revista de Direito e Tecnologia, Vol. 1, n.º 2, pp. 143-173.

PERRY, Walter L., MCINNIS, Brian, PRICE, Carter C., SMITH, Susan C. & HOLLYWOOD, John S., (2013). *«Predictive policing: the role of crime forecasting in law enforcement operations»*. RAND Corporation. http://www.jstor.org/stable/10.7249/j.ctt4cgdcz

PINHOL, Ana M. M. F., (2022). "A inteligência artificial e a justiça encontros e desencontros", Lisboa: Centro de Estudos Judiciários, pp. 41-56. https://cej.justica.gov.pt/LinkClick.aspx?fileticket=ttrgZIvgv7E%3D&portalid=30

RAPOSO, Vera L., (2022). "Ex machina: preliminary critical assessment of the European Draft Act on artificial intelligence", *International Journal of Law and Information Technology*, Vol. 30, Issue 1, pp. 88–109, https://doi.org/10.1093/ijlit/eaac007

ROCHA, Manuel L. & PEREIRA, Rui S. (Coords.), (2020). «Inteligência Artificial & Direito», Coimbra: Almedina:

- MEDEIROS, João, (2020). "Inteligência Artificial e contencioso penal I", pp. 115-119.

· GONÇALVES, José R., (2020). "Inteligência Artificial e contencioso penal II", pp.121-123.
· ALVES, Paulo F., (2020). "Inteligência Artificial e gestão de grandes processos", pp. 189-197.

RODRIGUES, Anabela M. (Coord.), (2020). «A Inteligência Artificial no Direito Penal», Coimbra: Almedina:

· RODRIGUES, Anabela M., (2020a). "A justiça preditiva entre a americanização e a europeização", pp. 11-58.
· SOUSA, Susana A. de, (2020a). "Não fui eu, foi a máquina: teoria do crime, responsabilidade e inteligência artificial", pp. 59-93.
· FIDALGO, Sónia, (2020). "A utilização de inteligência artificial no âmbito da prova digital – direito fundamentais (ainda mais) em perigo", pp. 129-161.
· COSTA, João M. & ABRANTES, António M., (2020). "Os desafios da Inteligência Artificial da Perspetiva Transnacional: A Jurisdição e a Cooperação Judiciária", pp. 163-217.
· CARIA, Rui, (2020). "O caso *State v. Loomis* – a pessoa e a máquina na decisão judicial", pp. 245-265.

RODRIGUES, Anabela M., (2020b). "Medida da pena de prisão – desafios na era da inteligência artificial", *Revista de Legislação e de Jurisprudência*, Vol. 149, n.º 4021, pp. 258 – 272.

RODRIGUES, Anabela M., (2020c). "A Política Criminal no Estado de Direito do Século XXI – os desafios da segurança", *Revista Brasileira de Ciências Policiais*, Vol. 11, N.º 1, pp. 19-40.

RUSSELL, Stuart & NORVIG, Peter, (2013). «Inteligência Artificial», 3ª ed., trad. Regina Célia Simille de Macedo, Rio de Janeiro: Campus Elsevier, 2013.

SACHOULIDOU, Athina. (2023). "Going beyond the "common suspects": to be presumed innocent in the era of algorithms, big data and artificial intelligence". *Artificial Intelligence and Law*, 31. https://doi.org/10.1007/s10506-023-09347-w

SALES, Ana D. R.; COUTINHO, Carlos M. C. & PARAISO, Letícia V., (2021). "Inteligência Artificial e decisão judicial: (im)possibilidade do uso de máquinas no processo de tomada de decisão". *Revista de Processo, Jurisdição e Efetividade da Justiça*. Vol. 7, n.º 1, pp. 34 – 54.

SANTOS, Hugo L. dos, (2022). "Inteligência Artificial e Processo Penal". Nova Causa.

SCHWAB, Klaus, (2016). A Quarta Revolução Industrial. Trad. Daniel Moreira Miranda. São Paulo: Edipro.

SEARLE, John R., (1980). "Minds, Brains and Programs". *The Behavioral and Brain Sciences*, 3, Cambridge University Press, pp. 417-457.

SILVA, Eva S. M. da & FREITAS, Pedro M. (Coords.), (2022). «Inteligência Artificial E Robótica: Desafios para o Direito do Século XXI» 1.ª Ed., Gestlegal. https://hdl.handle.net/1822/80752

- GONÇALVES, Marco C., (2022). "Inteligência artificial e processo judicial: em busca da celeridade, da eficiência e da qualidade da justiça", pp. 269-286.
- BORGES, Patrícia S., (2022). "A utilização de inteligência artificial na justiça aos olhos da proposta de regulamento da União Europeia", pp. 287-304.

SOUSA, Susana A. de, (2020b). "Um Direito Penal desafiado pelo desenvolvimento tecnológico: alguns exemplos a partir das neurociências e da inteligência artificial". *Revista da Defensoria Pública da União,* n.º 14, pp. 21-37.

STARR, Sonja B., (2013). "Evidence-Based Sentencing and the Scientific Rationalization of Discrimination". *Stanford Law Review*, Forthcoming, U of Michigan Law & Econ Research Paper, n.º 13-014. https://ssrn.com/abstract=2318940

TURING, Alan M., (1950). "Computing Machinery and Intelligence", *Mind, A Quarterly Review of Psychology and Philosophy*, Vol. LIX, n.º 236, Oxford University Press, pp. 433-460. https://doi.org/10.1093/mind/LIX.236.433

ZAVRŠNIK, Aleš, (2021). "Algorithmic justice: Algorithms and big data in criminal justice settings". *European Journal of Criminology*, Vol. 18(5), pp. 623–642.

ZEDNER, Lucia, (2007). "Pre-crime and post-criminology?". *Theoretical criminology,* 11.2, pp. 261-281.

DOCUMENTOS

AMA, Guia para a Inteligência Artificial, 2020. Disponível em: https://tic.gov.pt/areas-tematicas/inteligencia-artificial

P9_TA(2023)0236 – Alterações aprovadas pelo Parlamento Europeu, em 14 de junho de 2023, 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". Disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_PT.html

COM (2021) 206 final. Bruxelas, 21.4.2021 – *"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".*

(2020/2016(INI)) – Resolução do Parlamento Europeu, de 6 de outubro de 2021, sobre *"A inteligência artificial no direito penal e a sua utilização pelas autoridades policiais e judiciárias em casos penais".* Disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2021-0405_PT.pdf

COM (2020) 65 final. Bruxelas, 19.02.2020 – *"Livro Branco sobre a inteligência artificial – Uma abordagem europeia virada para a excelência e a confiança".*

COM (2020) 64 final. Bruxelas, 19.2.2020 – Relatório da Comissão ao Parlamento Europeu, ao Conselho e ao Comité Económico e Social Europeu sobre *"As implicações em matéria de segurança e de responsabilidade decorrentes da inteligência artificial, da Internet das coisas e da robótica".*

CAHAI (2020) 23, Council of Europe. Feasibility Study, Ad hoc Committee on Artificial Intelligence.

European Union Agency for Fundamental Rights (FRA), 2020 – *Getting the future right. Artificial intelligence and fundamental rights.* Disponível em: https://fra.europa.eu/en/publication/2020/artificial-intelligence-and-fundamental-rights

Grupo de Peritos de Alto Nível sobre a Inteligência Artificial, (2019) – *Orientações éticas para uma IA de confiança*

INTERPOL/UNICRI, (2019) – "*Artificial Intelligence and Robotics for Law Enforcement"*

COM (2018) 237 final. Bruxelas, 25.04.2018 – *Inteligência Artificial para a Europa*

PÁGINAS DE INTERNET

CAMPOS, Ricardo (2023). "Autodeterminação informacional 4.0 e o tratamento de dados pelo Estado". *Revista Consultor Jurídico*. https://www.conjur.com.br/2023-fev-28/direito-digital-tratamento-dados-estado-limites-constitucionais#_ftn11 (consultado a 05/07/2023)

CEPEJ (2018) 14 "Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente", Estrasburgo, 3 de dezembro de 2018. https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0 (consultado a 23/06/2023)

Fair Trials (2022). "AI Act: EU must ban predictive AI systems in policing and criminal justice". Disponível em: https://www.fairtrials.org/articles/news/ai-act-eu-must-ban-predictive-ai-systems-in-policing-and-criminal-justice/ (consultado a 05/07/2023).

COMMENT ON: 881 N.W.2d 749 (Wis. 2016) *State v. Loomis*: Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessment in Sentencing. 130 *Harvard Law Review* 1530 (2017, March). https://harvardlawreview.org/print/vol-130/state-v-loomis/ (consultado a 15/06/2023)

Comunicado de Imprensa n.º 18/2023 do Tribunal Constitucional Federal sobre a Sentença de 16 de fevereiro de 2023. Disponível em: https://www.bundesverfassungsgericht.de/SharedDocs/Pressemitteilungen/EN/2023/bvg23-018.html (consultado a 05/07/2023)

"Portugal é o 5.º país da UE com maior duração dos processos nos tribunais". *TSF*, 8 de junho de 2023. https://www.tsf.pt/portugal/sociedade/portugal-e-o-5-pais-da-ue-com-maior-duracao-dos-processos-nos-tribunais-16496519.html (consultado a 23/06/2023).

FERREIRA, Filipa & CAPUCHO, Luísa (2023). "Inteligência Artificial: "A tecnologia é imparável", mas precisa de "regulamentação adequada"". *JPN*. https://www.jpn.up.pt/2023/06/16/inteligencia-artificial-a-tecnologia-e-imparavel-mas-precisa-de-regulamentacao-adequada/ (consultado em 14/07/2023).

SILVA, Mário T. da. (2022). "Policiamento preditivo e direitos fundamentais". *Jornal Expresso*. Disponível em https://expresso.pt/opiniao/2022-12-30-Policiamento-preditivo-e-direitos-fundamentais-46fce25a (consultado a 23/06/2023).

SILVA, Susana C. e, (2021). "A Inteligência Artificial fraca e a fraca Inteligência Artificial", *Jornal Económico*. https://jornaleconomico.pt/noticias/a-inteligencia-artificial-fraca-e-a-fraca-inteligencia-artificial-764118/ (consultado a 23/06/2023).

Porto Editora – *silogismo* no Dicionário Infopédia da língua portuguesa. Porto: Porto Editora. Disponível em https://www.infopedia.pt/dicionarios/lingua-portuguesa/silogismo (consultado a 29/06/2023).

"ChatGPT: Everything you need to know about OpenAI's GPT-4 tool". https://www.sciencefocus.com/future-technology/gpt-3/ (consultado a 14/07/2023).

"DALL-E: a aplicação de Inteligência Artificial que cria imagens a partir de qualquer texto" Disponível em: https://shifter.pt/2021/01/dall-e-gpt3-openai/ (consultado a 14/07/2023).

Estratégia Nacional De Inteligência Artificial. Disponível em: https://www.incode2030.gov.pt/aip-2030/ (consultado a 14/07/2023).

Transformação Digital da Justiça 2015-2022. Disponível em: https://justica.gov.pt/Transformacao-Digital-da-Justica-2015-2022 (consultado a 14/07/2023).