

COLÓQUIOS

DO SUPREMO TRIBUNAL DE JUSTIÇA

TRIBUNAIS E INTELIGÊNCIA ARTIFICIAL
UMA ODISSEIA NO SÉCULO XXI

11 MAIO 2023



SUPREMO
TRIBUNAL
DE JUSTIÇA

COLEÇÃO LIVROS DIGITAIS DO SUPREMO TRIBUNAL DE JUSTIÇA

Esta coleção tem como objectivo principal coligir, de forma sistemática, os textos que resultem das comunicações levadas a efeito no Ciclo de Colóquios do Supremo Tribunal de Justiça, atividade sistémica inserida no Plano de Atividades da atual presidência do Supremo Tribunal de Justiça. Tem por escopo disponibilizar, a toda a comunidade jurídica, o acesso livre e gratuito dos conteúdos dos colóquios, de uma forma universal, potenciada pelo modo de divulgação/disponibilização digital.

Presidente do Supremo Tribunal de Justiça
Juiz Conselheiro Henrique Araújo

Chefe do Gabinete do Presidente do Supremo Tribunal de Justiça
Juíza Desembargadora Gabriela Cunha Rodrigues

Ficha Técnica

Coordenação Científica e Institucional do Colóquio

**Coordenação Institucional do Estudo sobre
“Automação, Digitalização e Inteligência Artificial nos Tribunais Portugueses”**
João da Silva Miguel, Juiz Conselheiro do Supremo Tribunal de Justiça, Jubilado
Sandra dos Reis Luís, Juíza de Direito e Adjunta do Gabinete do Presidente

Produção Executiva e revisão final do Livro Digital

Sandra dos Reis Luís, Juiz de Direito e Adjunta do Gabinete do Presidente

Coordenação Executiva do Colóquio

Gabinete do Presidente do Supremo Tribunal de Justiça

Fotografia

Ana Coelho, técnica Especialista do Gabinete do Presidente

Grafismo

Ana Oliveira Pinto, Designer

Esta publicação não adopta o novo Acordo Ortográfico, deixando-se essa opção ao critério dos autores.

Edição Dezembro de 2023

ISBN
978-989-53058-8-9

Sempre que desejar voltar ao índice, clique



07 **Henrique Araújo**
Juiz Conselheiro, Presidente do Supremo Tribunal de Justiça

09 **DISCURSO DE ABERTURA**
Henrique Araújo
Juiz Conselheiro, Presidente do Supremo Tribunal de Justiça

12 **NOTA CIENTÍFICA**
João Silva Miguel
Juiz Conselheiro Jubilado do Supremo Tribunal de Justiça

16 **A OPORTUNIDADE DO COLÓQUIO**
Sandra dos Reis Luís
Juíza de Direito, Adjunta do Gabinete do Presidente
Membro da Comissão de Coordenação e Acompanhamento
da Cimeira Judicial Ibero-Americana

CONFERÊNCIA DE ABERTURA E DEMAIS INTERVENÇÕES

20 **DIGITALIZAÇÃO, AUTOMAÇÃO, E INTELIGÊNCIA ARTIFICIAL NOS TRIBUNAIS JUDICIAIS PORTUGUESES**
DIGITALIZATION, AUTOMATION AND ARTIFICIAL INTELLIGENCE IN THE PORTUGUESE COURTS

Andreia Martinho
Investigadora no Data Intensive Studies Center (DISC), Tufts University

38 **COMO ESTÁ A INTELIGÊNCIA ARTIFICIAL A REVOLUCIONAR O DIREITO: UM CAMINHO SEM RETORNO**

Paulo Novais
Professor Catedrático da Universidade do Minho, ISLab, Algoritmi Center, LASI

48 **APLICAÇÕES DE GRANDES MODELOS DE LINGUAGEM NA ÁREA JURÍDICA**

Arlindo Oliveira
INESC-ID / Instituto Superior Técnico

54 **JUIZ IA: FUTURO, UTOPIA OU DISTOPIA?**

Vera Lúcia Raposo
Professora Auxiliar, NOVA School of Law

68 **BREVES REFLEXÕES SOBRE ÉTICA E A SUA APLICAÇÃO À INTELIGÊNCIA ARTIFICIAL**

Nuno Lima da Luz
Advogado Associado Sénior, Cuatrecasas

TROCA DE EXPERIÊNCIAS COM PAÍSES INTEGRANTES DA CIMEIRA JUDICIAL IBERO-AMERICANA

86 **INTELENCIA ARTIFICIAL (I.A) Y SISTEMAS DE JUSTICIA: TRES TEMÁTICAS Y ALGUNAS APLICACIONES CONCRETAS EN EL PODER JUDICIAL DE ESPAÑA**

Juan Martínez Moyaz
Magistrado y Vocal del consejo general del Poder judicial de España
Miembro de la Comisión de Coordinación y Seguimiento de la Cumbre Judicial Iberoamericana



98 LA APLICACIÓN ETICA A LA INTELIGENCIA ARTIFICIAL EN LAS RELACIONES JURIDICAS

Luís Maria Benitez Riera
 Ministro da Corte Suprema, Paraguai
 Miembro de la Comisión de Coordinación y Seguimiento de la Cumbre Judicial Iberoamericana

102 PODER JUDICIAL DE PUERTO RICO: DOS INICIATIVAS BASADAS EN INTELIGENCIA ARTIFICIAL

Sigfrido Steidel Figueroa
 Director Administrativo de los Tribunales de Puerto Rico y Juez del Tribunal de Apelaciones
 Miembro de la Comisión de Coordinación y Seguimiento de la Cumbre Judicial Iberoamericana

106 A INTELIGÊNCIA ARTIFICIAL E O PROJETO IRIS/SAMA NO SUPREMO TRIBUNAL DE JUSTIÇA

Gabriela Cunha Rodrigues
 Juíza Desembargadora
 Chefe do Gabinete do Presidente do Supremo Tribunal de Justiça

110 ES NECESARIA LA ELABORACIÓN DE PAUTAS PARA EL DESAROLLO DE LOS SISTEMAS DE INTELIGENCIA ARTIFICIAL EN MATERIA JUDICIAL?

John Perez Brignani
 Ministro da Corte Suprema do Uruguai
 Miembro de la Comisión de Coordinación y Seguimiento de la Cumbre Judicial Iberoamericana

116 ENCERRAMENTO

Agostinho Torres
 Juiz Conselheiro do Supremo Tribunal de Justiça

À MARGEM DO COLÓQUIO

130 IMPULSO Y CONSTITUCIONALIZACIÓN DE LOS NUEVOS DERECHOS FUNDAMENTALES

John Perez Brignani
 Ministro da Corte Suprema do Uruguai

144 MOMENTO MUSICAL

Clique para assistir ao Colóquio na integra.



Clique para assistir ao Debate do 1.º Painel.



Clique para assistir ao Debate do 2.º Painel.





VIRTUTIBVS
MAIORVM

VT. SIT. OMNIBVS. DOCUMENTO. P. P. D.



Henrique Araújo

Juiz Conselheiro, Presidente do Supremo Tribunal de Justiça

O Supremo Tribunal de Justiça tem realizado várias conferências e colóquios sobre os mais variados temas ligados ao Direito e à prática judiciária, com a participação de prestigiados juristas e de especialistas em matérias que de algum modo se conexas com a vida nos tribunais.

A valia jurídico-científica das comunicações produzidas impõe-nos o dever de encontrar os meios de divulgação mais expeditos e práticos para fazer chegar esses conhecimentos ao maior número possível de beneficiários.

Os livros digitais (*e-books*) são uma das ferramentas mais eficazes de divulgação de conteúdos, aliando a vantagem da praticidade ao insignificante encargo financeiro de edição, e permitindo ao beneficiário a consulta gratuita dos textos.

Por isso, esta forma de divulgação dos conteúdos das intervenções em conferências e colóquios passará a estar disponível na página oficial do Supremo Tribunal de Justiça, esperando-se que assim se atinja o objetivo de disponibilizar a toda a comunidade jurídica e aos cidadãos em geral a informação mais atualizada e completa sobre as matérias de interesse jurídico e judiciário.



DISCURSO DE ABERTURA

Henrique Araújo

Juiz Conselheiro, Presidente do Supremo Tribunal de Justiça

Clique para assistir.



- Senhora e Senhor Vice-Presidentes do Supremo Tribunal de Justiça
- Senhor Vice-Presidente do Conselho Superior da Magistratura
- Senhor Dr. Duarte Geraldês, em representação da Senhora Provedora de Justiça
- Senhora Bastonária da Ordem dos Advogados
- Senhora Dra. Susana Antas Videira, em representação do Senhor Bastonário da Ordem dos Solicitadores e Agentes de Execução.
- Senhoras e Senhores Conselheiros
- Senhoras e Senhores Procuradores-Gerais-Adjuntos
- Ilustres Convidados e Participantes
- Minhas Senhoras e Meus Senhores

É para mim um prazer e uma honra dar as boas-vindas aos participantes neste Colóquio, que conta também com a presença de Presidentes dos Supremos Tribunais e de Juizes Conselheiros e Coordenadores Nacionais dos países que compõem a Comissão de Coordenação e Acompanhamento da Cimeira Judicial Ibero-Americana, a quem dirijo uma especial saudação.

Para reflexão e debate sobre este tema tão entusiasmante, contamos com os reputados Oradores Andreia Martinho, Paulo Novais, Arlindo Oliveira, Vera Lúcia Raposo e Nuno Lima da Luz, sendo os dois painéis presididos pelos Juizes Conselheiros Maria Olinda Garcia e João Cura Mariano, e ficando o encerramento a cargo do Juiz Conselheiro Agostinho Torres.

A todos os meus sinceros agradecimentos.

A evolução tecnológica alcançou níveis que, poucos anos atrás, consideraríamos inatingíveis.

A recente ferramenta de Inteligência Artificial *ChatGPT* consegue programar, escrever artigos e até teses de mestrado. Bill Gates equipara-a ao aparecimento dos computadores pessoais nos anos oitenta.

A inteligência artificial tem-se estendido, a uma velocidade incrível, a vários domínios.

Veja-se o exemplo dos fármacos.

O processo de descoberta e desenvolvimento de fármacos, além do alto custo que lhe estava associado, demorava cerca de vinte anos.

Hoje, nesta área da ciência, têm sido implementadas técnicas de *machine learning* que consistem na utilização de redes neuronais artificiais para a criação de modelos inteligentes que mimetizam a capacidade de aprendizagem dos modelos biológicos.

Este é só um exemplo do uso da inteligência artificial numa área específica, mas muitos outros poderiam ser descritos em diversas atividades.

Como reagirá a Justiça à crescente intromissão desta e doutras ferramentas tecnológicas nas nossas vidas? Poderá incorporá-las na atividade dos tribunais? Em que medida e quais os perigos?

As análises sobre a era digital e as novíssimas tecnologias vão oscilando entre o revolucionário e o conservador, o entusiasmo e o ceticismo.

Antoine Garapon e Jean Lassège partilham uma visão muito pouco otimista sobre a matéria, bem espelhada na inquietação que formulam na seguinte pergunta: “Aceitariam ser julgados por algoritmos?”.

Em Portugal, como em alguns outros países, desmaterializaram-se os processos, adotaram-se meios avançados de assinatura digital, constituíram-se bases de dados e dotaram-se os tribunais de meios tecnológicos.

Na tramitação e consulta de dados, o uso de ferramentas de inteligência artificial poderá significar poupança de trabalho burocrático e libertará o juiz para funções mais substantivas.

Mas não se anteveem muitas outras aplicações da inteligência artificial na atividade dos tribunais.

De facto, sem prejuízo de perspectivas promissoras em algumas áreas, como no domínio dos *smart contracts* e do *blockchain*, a capacidade da máquina para reproduzir competências semelhantes às humanas não pode substituir o trabalho humano nas situações complexas da vida, como é o caso do ato de julgar.

A inteligência que comanda o processo de decisão judicial pertencerá sempre aos juízes.

A análise probatória e as diversas operações de raciocínio jurídico-normativo da sentença estão afastadas da «apreciação» automática e não dispensam a intervenção da avaliação, da sensibilidade e do bom senso dos juízes.

No entanto, a inteligência artificial poderá ser um instrumento importante em etapas decisivas para a formação da decisão, nomeadamente as que carecem de operações de cálculo, de pesquisa de fontes, de despiste de contradições lógicas e

de procura de precedentes.

Mesmo nas hipóteses em que poderia supor-se o uso da inteligência artificial (que seriam apenas os casos que envolvessem reduzidíssima densidade jurisdicional), haveria sempre que assegurar-se a confiabilidade e transparência do processo de tomada de decisão algorítmica e o respeito pelos direitos fundamentais.

Não sendo de descuidar, claro está, o respetivo quadro ético da inteligência artificial.

Neste aspeto, o Supremo Tribunal de Justiça apresentou, no âmbito da XXI edição da Cimeira Judicial Ibero-Americana, um projeto que veio a dar origem, em parceria com outros dois países integrantes, ao projeto «A Inteligência Artificial e as suas aplicações na Perspetiva Judicial».

Este projeto foi aprovado, por unanimidade, na I Reunião Preparatória, que se realizou na Cidade do México nos dias 26 a 28 de abril de 2022 e conta, agora, com a participação de quinze países, que o desenvolverão com vista à apresentação de um Código de Ética para a utilização da inteligência artificial no processo judicial.

Na Comissão Europeia, desde abril de 2021 que se arrastam as discussões sobre a Proposta de Regulamento do Parlamento Europeu e do Conselho, onde estão em confronto duas tendências: uma mais severa e restritiva quanto à regulamentação do uso da inteligência artificial e outra mais suave que incentiva o seu desenvolvimento.

Com a premência, cada vez mais acentuada, de regulamentação desta matéria, prevê-se que até ao final do corrente ano se aprove a Lei de Inteligência Artificial.

Ao que se pensa, essa lei tentará abrir espaço à inovação, garantindo simultaneamente a segurança e os direitos dos utilizadores.

Neste caminho sem retorno da inteligência artificial, daremos conta, neste Colóquio, de um estudo sobre Digitalização, Automação e Inteligência Artificial nos Tribunais Judiciais Portugueses, liderado pela Professora Doutora Andreia Martinho, com base num inquérito dirigido pelo Supremo Tribunal de Justiça ao universo dos magistrados judiciais portugueses.

Falaremos dos grandes modelos de linguagem e navegaremos no labirinto da ética judicial, entre a inovação e a segurança jurídica.

Antes de terminar, queria deixar uma palavra de agradecimento à Senhora Juíza Sandra dos Reis Luís, adjunta do Gabinete da Presidência, pelo notável trabalho desenvolvido durante largos meses para a realização deste evento tão oportuno.

Espero que este Colóquio se revele útil e que corresponda às vossas expectativas. •

Henrique Araújo
LISBOA, 11 de maio de 2023



NOTA CIENTÍFICA



João Silva Miguel
Juiz Conselheiro do Supremo Tribunal de Justiça, Jubilado

A Senhora Juíza Sandra dos Reis Luís, adjunta do Gabinete da Presidência, que, como enalteceu o Senhor Conselheiro Presidente do Supremo Tribunal de Justiça, desenvolveu, laboriosamente, notável trabalho na organização do Colóquio do Supremo Tribunal de Justiça sobre *Tribunais e Inteligência Artificial: Uma Odisseia do Século XXI*, pediu-me umas breves palavras a abrir a publicação das intervenções então produzidas.

Interpretei-o como uma generosidade da sua parte, que deveria honrar, pela relevância e atualidade do tema, mais do que pelos meus conhecimentos sobre o mesmo ou participação nos trabalhos.

Em 1989, o Prof. Helder Coelho¹ dizia que «[n]os últimos anos a IA explodiu num sem número de aplicações de oportunidades, popularizou-se e ganhou um estatuto bem estabelecido nas comunidades científicas a nível internacional».

Mais de três décadas depois, a sua aplicação ao sistema de justiça, no nosso país e em muitos outros da Europa e do mundo, exhibe-se ainda numa fase debutante, qual alegoria do gigante Adamastor, que o título deste Colóquio parece induzir: uma viagem extraordinária, pelos receios, medos e mistérios do desconhecido.

Não são claros os contornos do conteúdo e limites da Inteligência Artificial (IA), sendo múltiplas as suas dimensões interpretativas.

Segundo uma definição, a inteligência artificial é uma teoria e desenvolvimento de sistemas computacionais capazes de realizar tarefas que historicamente exigem inteligência humana, como o reconhecimento da fala, a tomada de decisões e a iden-

¹ *Inteligência Artificial – O balanço da década de oitenta*, Ed. Caminho, 1991, pp. 11-12.

tificação de padrões, consistindo numa vasta gama de tecnologias, onde se inclui o *machine learning*, o *deep learning*, e o processamento da linguagem natural (NLP).

O recente Regulamento Inteligência Artificial afirma-a como uma família de tecnologias em rápida evolução que contribui para um vasto conjunto de benefícios económicos, ambientais e sociais em todo o leque de indústrias e atividades sociais, podendo contribuir para progressos sociais e ambientais, nomeadamente na justiça.

Apesar do exponencial e veloz desenvolvimento que tem conhecido, em 2018 o uso de algoritmos de inteligência artificial nos sistemas de justiça europeus restringia-se primordialmente ao sector privado com o objetivo de desenvolver soluções comerciais para seguradoras, departamentos jurídicos, advogados e particulares².

Nessa altura, alguns países – Países Baixos, Reino Unido, Letónia e Estónia – davam os primeiros passos na adoção de soluções de IA no quadro do contencioso cível de pequenas causas, com o propósito de assim resolver a acumulação desses casos.

Em França, no ano de 2017, foi implementado um projeto piloto em dois tribunais de relação sobre justiça preditiva⁴ (Plataforma PREDICTICE), sem resultados promissores⁴.

Noutro estádio de evolução, nos Estados Unidos da América, a aplicação do programa COMPAS, acrónimo de *Correctional Offender Management Profiling for Alternative Sanctions*, fazia o seu caminho.

A maior notoriedade adveio da condenação do arguido *Eric Loomis*, no caso *Loomis v. Wisconsin*, e do recurso interposto perante o Supremo Tribunal do Wisconsin, sob a alegação de que o *software* de avaliação de risco de código fechado usado na sua condenação violava o seu direito a um processo equitativo, por impedir-lo de contestar a validade científica e a precisão da fonte do algoritmo, além de levar em consideração questões de género e de raça.

Ao recurso foi negado provimento, em decisão unânime de 16 de julho de 2016, por ter sido entendido que, apesar de o relatório anexado ao processo não explicar como o Programa COMPAS utiliza informações para calcular as pontuações de risco, um guia divulgado ao público esclarece que as pontuações do Programa são baseadas em grande parte em informações estáticas do histórico criminal do arguido,

² *European Ethical Charter on the use of artificial intelligence (AI) in judicial systems and their environment*, preparada pela *European Commission for the Efficiency of Justice* (CEPEJ), Edição do CoE, 2019, p.16.

³ O conceito justiça preditiva é controverso, com ele pretendendo referir-se à análise de grande volume de decisões judiciais, utilizando tecnologias de inteligência artificial para construir, relativamente a certos tipos de litígios especializados, previsões quanto ao resultado do mérito. Sobre justiça preditiva, vd. Antoine Garapon e Jean Lassègue, *Justice digitale*, Puf, 2018, pp. 219-244.

⁴ Vd. *Themis Annual Journal*, 2019, Vol. I, Issue I, outubro de 2019, 1.ª ed., p. 243 e ss, trabalho da equipa CEJ, composta pelos auditores de justiça Elisa Alfaia Sampaio, João J. Seixas, Paulo Jorge Gomes, sendo tutora a desembargadora Patrícia Helena Costa) sobre *Inteligência artificial e decisão judicial* (Artificial Intelligence and the Judicial Ruling), também publicado no e-book sobre os trabalhos dos auditores de justiça do Concurso Themis 2019, CEJ, 2019, p. 63 e ss.

com uso limitado de algumas variáveis dinâmicas, como os antecedentes criminais (*criminal associates*) e o histórico do “abuso de substâncias” da pessoa visada⁵.

Na Europa, o Tribunal Europeu dos Direitos Humanos (TEDH) examinou dois casos, relevando do digital e da inteligência artificial, ambos apreciados no âmbito do direito a um processo equitativo, reconhecido no artigo 6.º da Convenção Europeia dos Direitos Humanos (CEDH).

No primeiro – *Marcello Viola c. Itália*, de 5 de outubro de 2006 –, o Tribunal concluiu que a participação, através de videoconferência, da pessoa em causa na audiência de recurso que lhe interessava, acompanhada de advogado, «não colocou a defesa em desvantagem substancial em comparação com as outras partes no processo», pois teve a oportunidade de exercer os direitos e prerrogativas inerentes ao conceito de um julgamento justo, conforme consagrado no artigo 6.º da Convenção Europeia; no segundo caso – Acórdão, de 4 de junho de 2019, proferido no caso *Sigurður Einarsson e Outros c. Islândia*, §92 – os requerentes queixavam-se de não terem tido um julgamento justo, por não terem tido acesso a todas as provas no processo, ao ser-lhes negada a possibilidade de pesquisar massas de dados eletrônicos (*full collection of data*) resultantes de buscas e apreensões, selecionados com base em programa usado (*Clearwell*) e respetivas palavras-chave, ao que o Tribunal respondeu que os requerentes não tendo fornecido qualquer especificação do tipo de provas que procuravam, ficou convencido de que a falta de acesso aos dados em questão não foi tal que os impedisse de ter um julgamento justo.

Apesar dos reconhecidos grandes avanços tecnológicos em matéria de inteligência artificial, os últimos anos parecem não ter desencadeado particulares respostas nos sistemas de justiça nacionais.

A última *Opinion* do Conselho Consultivo dos Juízes Europeus (CCJE) confirma-nos, com base em dados recentes dos países do Conselho da Europa considerados, que «a utilização da inteligência artificial (IA) (...) permanece numa fase inicial de desenvolvimento», estando a ser «ativamente explorada em alguns Estados», concentrada em «tarefas administrativas, onde se incluem os programas de conversão de voz em texto (ditado), a tradução de idiomas e o anonimato automático de sentenças e despachos»⁶.

Um melhor conhecimento e compreensão sobre esta tecnologia, suas capacidades, benefícios e riscos para a administração da justiça, apresenta-se como imperioso.

Por isso, este colóquio é de extrema relevância: pela atualidade e pertinência do tema e pela necessidade de trazer para o debate a reflexão sobre os desafios que esta realidade nos convoca, e que os magistrados e em particular os juízes serão confrontados.

⁵ Pedida a intervenção do Supremo Tribunal dos Estados Unidos da América, este não conheceu do recurso, em maio de 2017.

⁶ Vd. *Opinion* n.º 26(2023), de 1 de dezembro de 2023, do Conselho Consultivo de Juízes Europeus do Conselho da Europa, sob o tema *Moving forward: the use of assistive technology in the judiciary*, §5.

A abordagem de soluções ensaiadas noutros países, pela dimensão internacional do Colóquio, constitui um elemento de comparação e de confirmação do acerto dos passos dados e dos recursos alocados.

Do estudo empírico sobre Digitalização, Automação, e Inteligência Artificial nos Tribunais Judiciais Portugueses, realizado pela Investigadora Dra. Andreia Martinho, para apoio ao Colóquio, destacaria um traço essencial, reconhecido pelos participantes naquele: a formação de Juízes deve incluir uma maior componente de preparação sobre os desafios jurídicos associados às tecnologias de Inteligência Artificial.

Esta conclusão está em linha com as estratégias de formação judiciária, que foram delineadas para o período 2020-2024 e 2020-2027, pela Comissão Europeia e pela Rede Europeia de Formação Judiciária, onde expressamente se afirma que a formação judiciária deve preparar os profissionais da justiça para tirarem partido da digitalização e para utilizarem a inteligência artificial⁷.

É fundamental que assim seja, e que se avance mais.

Nas minhas responsabilidades na direção do Centro de Estudos Judiciários, entusiasmei-me com a ideia da criação de um centro de estudos de Inteligência Artificial, envolvendo, em especial, as magistraturas e a academia, para, de modo interdisciplinar e integrado, estudar a aplicação dessa tecnologia à realidade do sistema de justiça nacional.

Foram feitos contactos, de grata recetividade, mas a pandemia e a aproximação do termo do mandato não consentiram um melhor desenvolvimento da ideia e a sua concretização.

Continuo convicto que traria inegáveis contributos à Justiça e ao país, com o melhor estudo e conhecimento sobre as suas virtualidades e potencialidades, quanto à qualidade, confiança e celeridade que poderia oferecer à justiça, e aos riscos que encerra e forma de os acautelar, nomeadamente pela observância de princípios éticos sólidos comumente reconhecidos, além de contribuir para uma melhor literacia neste âmbito e perceber se os receios, medos e mistérios na odisséia a empreender se mostram fundados.

Pois, como disse ROBERT SPANO, anterior Presidente do Tribunal Europeu dos Direitos Humanos, «como Juízes, estamos todos sob uma certa pressão para um desempenho mais eficiente e para proferir justiça mais célere. A inteligência artificial oferece certas oportunidades em termos de processamento do caso. No entanto, é preciso compreender e gerir claramente os riscos para os direitos humanos»⁸. •

⁷ Como se previa na Estratégia de formação judiciária europeia para 2021-2024, e se harmonizava com a Estratégia para a Formação Judiciária 2021-2027 da Rede Europeia de Formação Judiciária. Na presidência portuguesa do Conselho da União Europeia (2021), a digitalização era mesmo um dos eixos basilares na área da justiça.

⁸ Discurso de abertura do Seminário *The Rule of Law and Justice in a digital age*, organizado no contexto do «Dialogue between Judges», que teve lugar em Estrasburgo, em 10 de setembro de 2021. Vd. «Dialogue between judges, European Court of Human Rights, Council of Europe, 2021», p. 5.

A OPORTUNIDADE DO COLÓQUIO



Sandra dos Reis Luís

Juíza de Direito

Adjunta do Gabinete do Presidente

Membro da Comissão de Coordenação e Acompanhamento da Cimeira Judicial Ibero-Americana

Iniciamos, há cerca de um ano, os trabalhos de investigação que permitiram traçar este projecto.

Entre esboços, avanços e retrocessos, a vontade ganhou corpo.

Muito graças a todos os intervenientes/oradores, aos colegas que integraram os grupos de teste do inquérito, aos que responderam ao mesmo, que permitiu a realização do estudo, cujas primeiras conclusões nos são trazidas pela investigadora Doutora Andreia Martinho.

Muito graças ao Senhor Juiz Conselheiro João Miguel, que não hesitou a responder ao desafio e foi uma companhia inexcelável em todo o percurso.

Graças, acima de tudo, a Sua Excelência, o Presidente do Supremo Tribunal de Justiça, Juiz Conselheiro Henrique Araújo, um homem de mente aberta aos desafios que o conhecimento nos traz.

Tivemos a intenção de, aproveitando a presença dos membros integrantes da Comissão de Coordenação e Acompanhamento da Cimeira Judicial Ibero-Americana, em território português, e, concretamente, no Supremo Tribunal de Justiça, proporcionar uma troca de experiências que nos pudesse dar a conhecer as práticas e avanços de cada um dos países em matéria de automação, digitalização e inteligência artificial.

A inteligência artificial – em contraposição ao que é humano – apresenta-se como uma realidade, cuja presença não podemos ignorar. Cujos benefícios, também para o trabalho na justiça, no judiciário, não podem ser escamoteados, embora enfrentados com os cuidados que os riscos que lhe estão associados nos exigem.

Estamos convictos que a formação numa área ainda desconhecida para muitos de nós, em constante, diria diária, evolução nos merece a devida atenção, a iniciar pela reflexão e prática de uma formação adequada, que deve iniciar-se nas faculdades e no Centro de Estudos Judiciários, mas a que o Supremo Tribunal de Justiça é bastante sensível.

Os grandes modelos de linguagem, a quantidade e a qualidade dos dados que alimentam as máquinas (a complexidade que a obrigação de anonimização introduz relativamente aos dados pessoais, o uso de dados anonimizados, de dados sintéticos), a nova geração de algoritmos, a busca pelo algoritmo perfeito, que Pedro Domingos, apelidou de “algoritmo mestre”, os vieses, os preconceitos de quem programa, a interdisciplinaridade dos vários intervenientes na introdução de dados, os princípios enformadores, designadamente éticos, a monitorização, a certificação, estes cuja presença em todo o processo é essencial para garantir o respeito pelos direitos humanos, a responsabilidade, são questões de grande relevância e sofisticação, cujo desenvolvimento todos devemos acompanhar a passos largos.

Com este Colóquio “Tribunais e Inteligência Artificial – uma odisséia no século XXI”, pretendeu-se, sobretudo, alertar para o advento da IA, para a utilidade que a mesma pode representar no trabalho que se desenvolve nos tribunais, para imprimir celeridade de resposta sem olvidar a qualidade da mesma e alertar para os riscos e para a necessidade de aprendizagem, que a todos deve envolver, numa perspectiva tão aberta, quanto crítica e cautelosa.

Termina-se acentuando que, não obstante todas as vantagens que a IA nos trará, existem características essenciais ao julgamento e ao julgador (também aos mandatários das partes *lato sensu*) que, por serem intrinsecamente humanas, não poderão – sobretudo, em questões complexas – ser apreendidas pelas máquinas, como certamente concluirão através da leitura dos vários artigos que agora se divulgam com o presente Livro Digital.

As máquinas – a IA estreita – por lhes faltarem capacidade empática, de improviso, intuitiva, flexibilidade, os tais Juízes IA não substituirão – pelo menos, no actual estado da arte – o Juiz humano. •

CONFERÊNCIA
DE ABERTURA
E DE MAIS
INTERVENÇÕES

DIGITALIZAÇÃO, AUTOMAÇÃO, E INTELIGÊNCIA ARTIFICIAL NOS TRIBUNAIS JUDICIAIS PORTUGUESES



Andreia Martinho
Investigadora no Data Intensive Studies Center
(DISC), Tufts University

Clique para assistir.



RESUMO: Neste breve relatório apresentamos parte dos dados recolhidos no âmbito de estudo empírico sobre a Digitalização, Automação, e Inteligência Artificial nos Tribunais Judiciais Portugueses.

PALAVRAS-CHAVE: Inteligência Artificial, Justiça, Juízes, Tribunais Judiciais

1 | Introdução

A implementação de tecnologias de Digitalização, Automação, e Inteligência Artificial na Justiça apresenta um grande potencial para melhorar o acesso aos tribunais, a celeridade dos processos judiciais, e a justeza das decisões judiciais.

Estas tecnologias operam em três níveis distintos: (i) suporte, (ii) substituição, e (iii) disrupção. Ao nível mais básico (suporte), visam apoiar os profissionais da Justiça em tarefas relacionadas com a gestão documental, pesquisa jurídica, ou redação de documentos jurídicos. No nível intermédio (substituição), estas tecnologias assumem parte das funções relacionadas com a gestão do tribunal (por exemplo, triagem e marcação de presenças), gestão do processo judicial (por exemplo, automação de análise preliminar dos articulados), ou até mesmo a resolução de disputas de baixa complexidade. Ao nível mais avançado (disrupção), que ora se situa no plano especulativo, estas tecnologias podem eventualmente levar a novas formas de Justiça, que assentem por exemplo, num maior policiamento preditivo.

Apesar do potencial destas tecnologias, existem também desafios de natureza técnica (ex. escassez de dados estruturados), jurídica (ex. compatibilidade da automação com os valores de Justiça), e ética (ex. transparência dos algoritmos) inerentes à sua implementação na Justiça. É assim essencial promover uma inovação tecnológica que seja inclusiva, transparente e responsável. Neste contexto de inovação, as comunidades Jurídicas e Tecnológicas devem trabalhar em conjunto por forma a identificar, debater, e investigar soluções para os desafios associados às novas tecnologias.

Este estudo visa contribuir para a inovação tecnológica inclusiva, transparente, e responsável na Justiça, através da realização de inquérito a Juízes dos Tribunais Judiciais Portugueses sobre as oportunidades, riscos, e desafios das tecnologias de Digitalização, Automação, e Inteligência Artificial. Neste breve relatório apresentamos parte dos dados recolhidos no âmbito deste Estudo, referentes às Prioridades na Digitalização e Automação do Processo Judicial; Formação de Juízes no novo Paradigma Tecnológico; e Tecnologias de Inteligência Artificial e Precedente Jurisprudencial.

2 | Metodologia

O inquérito foi desenvolvido pela Investigadora Principal deste estudo na plataforma Qualtrics, e posteriormente testado em três sessões piloto com Juízes de diferentes Tribunais Judiciais. Após receber aprovação do *Institute Review Board da Tufts University* sobre a versão final do inquérito, o mesmo foi distribuído por email aos Juízes em funções nos Tribunais Judiciais Portugueses através de *link* que esteve disponível entre finais do mês de Março de 2023 e finais do mês de Abril de 2023. Os resultados do inquérito foram analisados em Python.

3 | Caracterização da Amostra

Foram registados 279 inquéritos completos (\leq três respostas incompletas). Consideramos que esta amostra é diversificada e representativa da população de Juízes em Portugal, no que respeita ao género, fase da carreira, e tipo de tribunal. Há uma prevalência de Juízes do género feminino, Juízes a meio e fim de carreira, Juízes a exercer funções em tribunal de Primeira Instância, e Juízes a exercer funções no Distrito de Lisboa (Tabela 1 & Figuras 1 e 2).

TABLE 1: AMOSTRA DO ESTUDO

Variável	N
Género	
Feminino	158
Masculino	117
Carreira	
Início (<10 anos)	34
Meio (10-25 anos)	114
Fim (>25anos)	129
Tribunal	
Tribunal de Primeira Instância	167
Tribunal da Relação	78
Supremo Tribunal de Justiça	34
Distrito	
Guimarães	37
Porto	44
Coimbra	21
Lisboa	131
Évora	46

FIGURE 1: GÉNERO E FASE DA CARREIRA DOS PARTICIPANTES NO ESTUDO

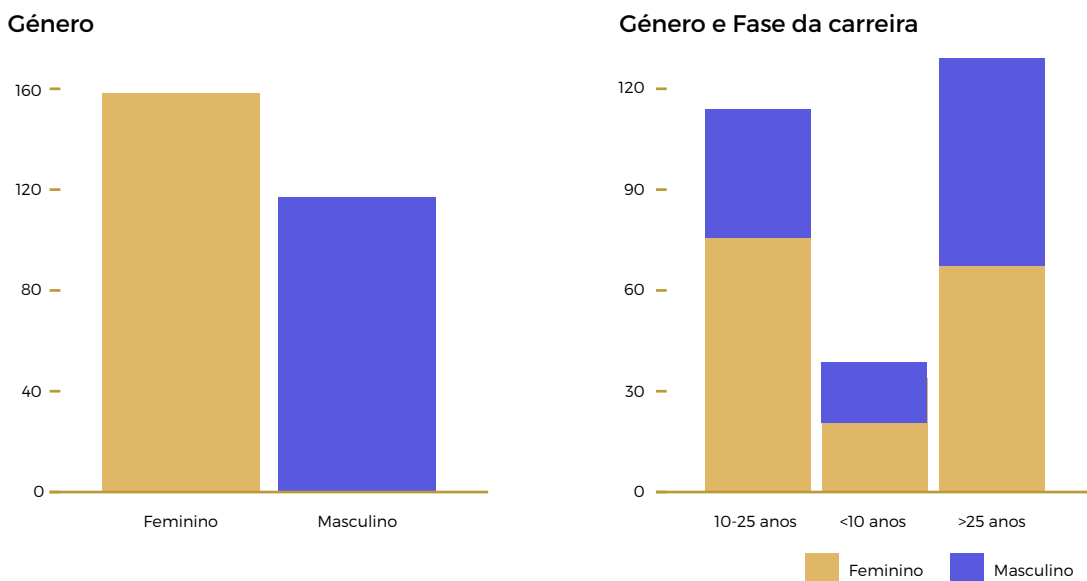
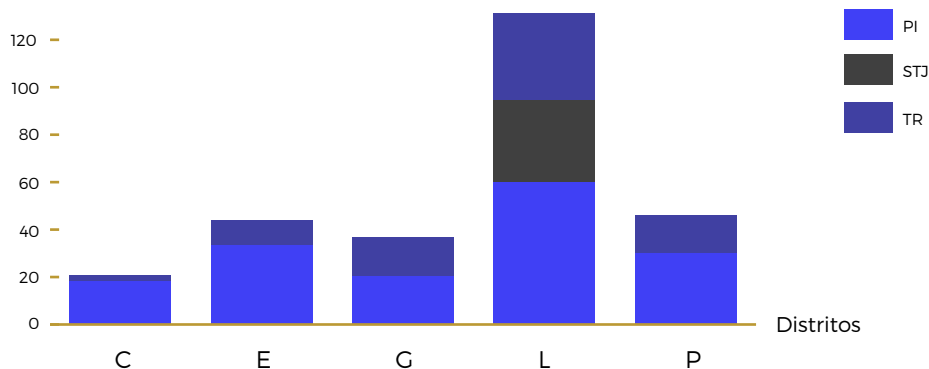


FIGURE 2: DISTRITO E TRIBUNAL ONDE OS PARTICIPANTES NO ESTUDO EXERCEM FUNÇÕES



4 | Resultados

4.1 Digitalização do Processo Judicial

Prioridades ao nível da Digitalização do Processo Judicial Os participantes neste Estudo consideram prioritário um sistema de gestão de processos que permita uma melhor visualização do processo, bem como aplicar um maior número de funcionalidades nos documentos (Tabela 2).

TABLE 2: PRIORIDADES NA DIGITALIZAÇÃO DO PROCESSO JUDICIAL

Prioridade	Sistema
1	Sistema de gestão de processos que permita uma melhor visualização do processo.
2	Sistema de gestão de processos que permita aplicar um maior número de funcionalidades nos documentos (por exemplo, editar documentos).
3	Digitalização integral dos processos, incluindo apensos.
4	Sistema de gestão de processos no qual a documentação é apresentada de forma dinâmica (por exemplo, hiperligações para diferentes documentos).
5	Outro*

**Outro* inclui soluções de digitalização sugeridas pelos participantes:

- Sistema de gestão dos suportes áudio e vídeo da documentação das audiências de julgamento.
- Despachos antecipados com assinatura agendada.
- Tablets que permitam o acesso ao histórico do processo, a cores e com aplicações úteis instaladas, como o Google Earth Pro.
- Sistema em que todos os operadores usem o mesmo processador de texto.
- Divulgação das decisões.
- Possibilidade de pesquisa por termos em todo o processo.
- Sistema de gestão de processos que permita um melhor controlo do fluxo de trabalho.
- Sistema de gestão de processos que seja rápido e não bloqueie constantemente.
- Documentos editável sem desconfiguração.
- Maior uso de templates com o preenchimento prévio dos intervenientes processuais.
- Sistema de transformação automática de áudio em texto editável.
- Sugestão de despachos pré-criados e editáveis, de mero expediente, de acordo com a fase processual.

- Peças processuais em formulário para os dados estruturados (factos, direito) poderem ser extraídos automaticamente.
- Correta designação de todos os atos mencionados no histórico do processo.
- Sistema que permita copiar textos e imagens para formato editável a usar pelo juiz.
- Programa que permita ao Juiz durante a diligência selecionar factos que possam ser relevantes para uma decisão.

4.2 Automação do Processo Judicial

Prioridades ao nível da Automação do Processo Judicial

Os participantes neste Estudo consideram prioritário sistemas de contagem de prazos, determinação de competência territorial, e cálculo de pensões de alimentos/juros de mora/fundos de garantia/atualizações (Tabela 3).

TABLE 3: PRIORIDADES NA AUTOMAÇÃO DO PROCESSO JUDICIAL

Prioridade	Sistema
1	Contagem de prazos (por exemplo, prazo processual, liquidacção de pena)
2	Determinacção da competência territorial
3	Cálculo de pensões de alimentos/juros de mora/fundos de garantia/actualizações
4	Determinação de custas judiciais
5	Cálculo de indemnizações
6	Determinação do valor da acção
7	Fixação do cúmulo jurídico
8	Outro*

**Outro* inclui soluções de automação sugeridas pelos participantes:

- Indicação automática de Doutrina e Jurisprudência para a resolução do caso.
- Transcrição em texto editável do certificado de registo criminal do arguido.
- Automatização de prazos legalmente definidos.
- Indicação do prazo previsível de duração do processo em dias, atualizável em função das incidências processuais que se vão sucedendo.
- Elaboração de projeto de decisão.

4.3 Intervenção Tecnológica na Justiça

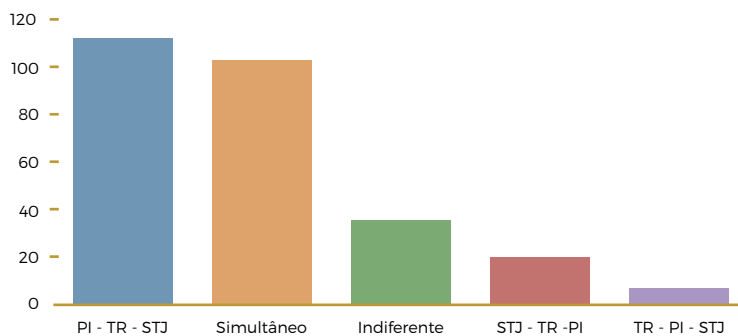
Processos mais Favoráveis a uma Intervenção Tecnológica ao nível da Digitalização, Automação, e Inteligência Artificial

Os participantes neste Estudo consideram que o Processo Civil é o processo mais favorável a uma intervenção tecnológica e, pelo contrário, o Processo Criminal é o menos favorável a essa intervenção (Tabela 4). No que respeita à ordem mais adequada, a maioria dos participantes consideram que a inovação tecnológica deve começar pelos Tribunais de Primeira Instância, seguido dos Tribunais da Relação e por fim do Supremo Tribunal de Justiça ou, em alternativa, em todos os tribunais em simultâneo.

TABLE 4: PROCESSOS JUDICIAIS MAIS FAVORÁVEIS A UMA INTERVENÇÃO TECNOLÓGICA AO NÍVEL DA DIGITALIZAÇÃO, AUTOMAÇÃO, E INTELIGÊNCIA ARTIFICIAL

Ranking	Processo
1	Civil
2	Social/Trabalho
3	Criminal

FIGURE 3: ORDEM ADEQUADA À INTERVENÇÃO TECNOLÓGICA AO NÍVEL DA DIGITALIZAÇÃO, AUTOMAÇÃO E INTELIGÊNCIA ARTIFICIAL NOS TRIBUNAIS JUDICIAIS.



4.4 Formação de Juízes no Novo Paradigma Tecnológico

Formação de Juízes

Os participantes neste estudo consideram que a formação de Juízes deve incluir uma maior componente de preparação sobre os desafios jurídicos associados às tecnologias de Inteligência Artificial. Estes resultados são consistentes quando isolamos os participantes em início de carreira (Figura 4). No que respeita ao momento em que essa formação deve ser oferecida, parece haver uma ligeira preferência para que a formação seja oferecida de

forma transversal, ou seja, na Licenciatura e/ou Mestrado em Direito (Faculdade de Direito = FD), Formação inicial de magistrados (FI) assegurada pelo Centro de Estudos Judiciários, e Formação contínua de magistrados (FC) assegurada pelo Centro de Estudos Judiciários (Figura 5).

FIGURE 4: FORMAÇÃO DE JUÍZES NO NOVO PARADIGMA TECNOLÓGICO

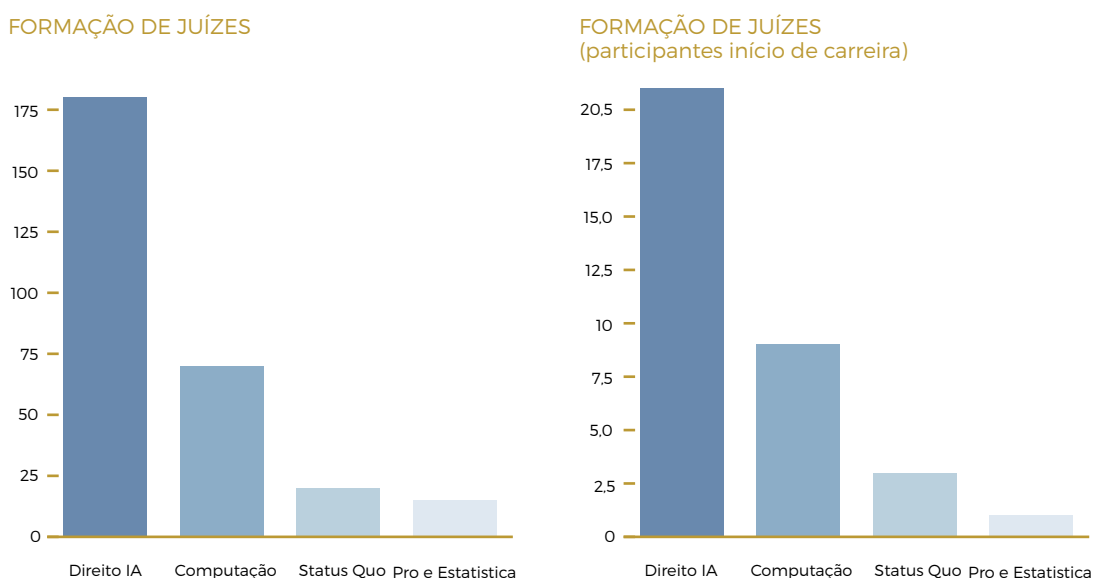
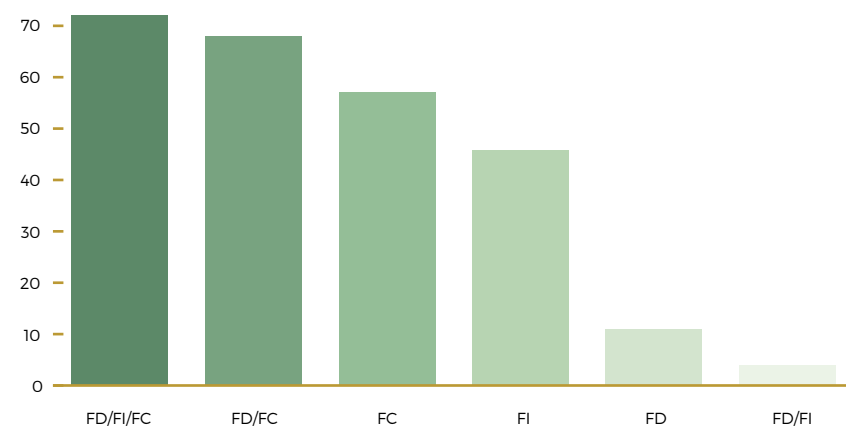


FIGURE 5: FASE DE FORMAÇÃO DE JUÍZES

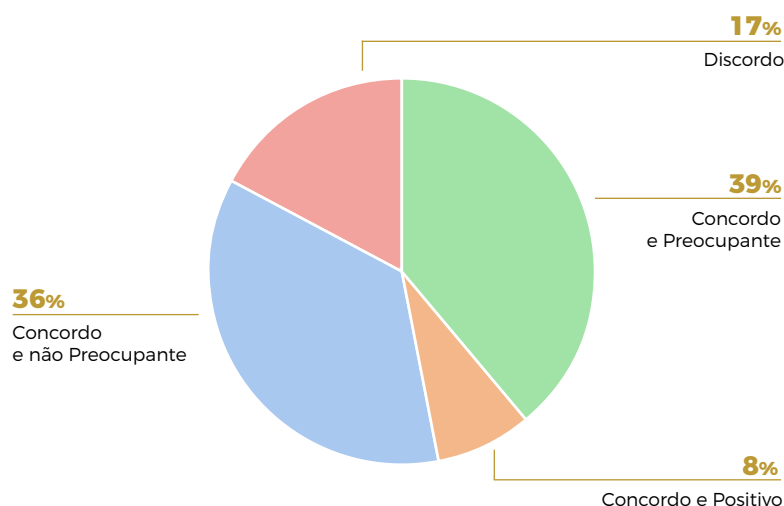


4.5 Precedente Jurisprudencial

Tecnologias de Inteligência Artificial e Precedente Jurisprudencial

Os participantes neste estudo consideram que, tendo em conta que as tecnologias de Inteligência Artificial identificam padrões nos dados para assim formular previsões, quando aplicadas no espaço jurídico, estas tecnologias podem levar a uma transição para um modelo jurídico mais próximo da Common Law, assentando em larga medida no precedente jurisprudencial (75%). Contudo, apenas aproximadamente metade desses participantes (39%) consideram esse cenário preocupante (Figura 6).

FIGURE 6: TRANSIÇÃO PARA UM MODELO JURÍDICO DE PRECEDENTE JURISPRUDENCIAL



5 | Discussão

Neste estudo investigámos as perspetivas dos Juízes dos Tribunais Judiciais sobre as oportunidades, riscos, e desafios das tecnologias de Digitalização, Automação, e Inteligência Artificial na Justiça. Consideramos que os resultados ora reportados apresentam grande relevância no atual contexto de transição tecnológica. Estes resultados devem assim ser tidos em consideração pelas comunidades Jurídicas, Tecnológicas, e Políticas envolvidas nos processos de inovação tecnológica na Justiça.

Quanto às Prioridades na Digitalização e Automação do Processo Judicial, consideramos, antes de mais, que os sistemas considerados prioritários devem merecer urgente atenção por parte das entidades responsáveis pela

gestão das tecnologias nos tribunais. Reconhecemos que a implementação alguns destes sistemas requer porventura um elevado esforço logístico e de investimento (por exemplo, digitalização integral dos processos, incluindo apensos), mas, pelo contrário, outros sistemas parecem ser de implementação bastante simples (por exemplo, a correta designação de todos os atos mencionados no histórico do processo, contagem de prazos e determinação de competência territorial). Algumas das soluções propostas pelos Juízes, por exemplo no que respeita a variados cálculos, desde a contagem de prazos ao cálculo de pensões e indemnizações, podem ser desenvolvidas em open source no âmbito da Formação Inicial e Contínua de Magistrados.

Sobre a Formação de Juízes no Novo Paradigma Tecnológico, os resultados deste estudo indicam que os Juízes consideram que a sua formação deve incluir uma maior componente de preparação sobre os desafios jurídicos associados às tecnologias de Inteligência Artificial. Assim sendo, é necessário refletir sobre as diferentes abordagens ao serviço da literacia jurídica sobre Inteligência Artificial. Ao invés de apresentar estas tecnologias como entidades externas, pensamos que devem ser devidamente enquadradas e harmonizadas no espaço jurídico. Por outro lado, anotamos que os participantes neste Estudo, não têm interesse em ter formação em Probabilidade e Estatística. Pese embora se compreenda que os Juízes sintam uma maior necessidade em ganhar conhecimentos que possam aplicar na resolução de disputas no Tribunal, ousamos nesta sede salientar a relevância da Probabilidade e Estatística como elemento crítico no paradigma atual de Inteligência Artificial de Dados. Como Ciência Aplicada, o Direito visa dotar os seus formandos de ferramentas que lhes permitam resolver problemas. Ao incluir fundamentos matemáticos e algorítmicos na formação de juristas, estes estarão mais bem preparados para compreender, acompanhar e desenvolver soluções inovadoras para problemas de natureza jurídica.

Neste breve relatório exploramos também as Tecnologias de Inteligência Artificial e Precedente Jurisprudencial. Estas tecnologias podem trazer mudanças importantes ao atual modelo jurídico, nomeadamente, segundo os participantes neste Estudo, no maior uso do precedente jurisprudencial. Pensamos que são necessários mais estudos para desenvolver as bases teóricas do modelo jurídico do futuro, assente numa maior componente tecnológica.

Por fim, cumpre mencionar que este Estudo reflete apenas as visões dos Juízes, e que, para que a inovação tecnológica seja verdadeiramente inclusiva, é necessário ouvir os outros profissionais da Justiça.

6 | Agradecimentos

A autora deste estudo faz um especial agradecimento aos Juízes Sandra dos Reis Luís, João Ferreira, e João Miguel Mendes, pela colaboração no desenvolvimento do inquérito. A autora agradece aos Juízes Sandra dos Reis Luís e João Miguel Mendes pela divulgação do inquérito. A autora agradece também aos Juízes que participaram nas sessões piloto: Ana Chambel, Cátia Santos, Filipe Aveiro Marques, Gabriela Cunha Rodrigues, Georgina Rodrigues, João Cura Mariano, Maria Olinda Garcia, e Pedro Soares de Albergaria. A autora agradece ao Supremo Tribunal de Justiça, na pessoa do seu Presidente, Juiz Conselheiro Henrique Araújo pela colaboração neste projecto de investigação. Por fim, a autora agradece a todos os Juízes que completaram o inquérito. •



DIGITALIZATION, AUTOMATION AND ARTIFICIAL INTELLIGENCE IN THE PORTUGUESE COURTS

Andreia Martinho
Tufts University

ABSTRACT: In this short report we present part of the data collected within the context of an empirical study on Digitization, Automation, and Artificial Intelligence in the Portuguese Judicial Courts.

KEYWORDS: Artificial Intelligence, Justice, Judges, Judicial Courts.

1 | Introduction

The implementation of Digitalization, Automation, and Artificial Intelligence technologies in the Judicial System holds great potential to improve access to courts, celerity in judicial adjudication, and fairness of judicial decisions.

These technologies operate at three distinct levels: (i) *Supportive*, (ii) *Replacement*, and (iii) *Disruption* [5]. At the most basic level (*supportive technologies*), they aim to assist legal professionals in tasks related to document management, legal research, or drafting legal documents [5, 1]. At the intermediate level (*replacement technologies*), these technologies take on some functions related to court management (e.g. triage and scheduling), judicial adjudication management (e.g. automation of preliminary analysis of pleadings), or even the resolution of low-complexity disputes [5, 4]. At the most advanced level (*disruptive technologies*), which is for now speculative, these technologies may eventually lead to new forms of Justice, such as predictive policing [5].

Despite the potential of these technologies, there are also technical (e.g. scarcity of structured data), legal (e.g. compatibility of automation with Justice values), and ethical challenges (e.g. transparency of algorithms) associated with their implementation in judicial systems. It is critical to promote technological innovation that is inclusive, transparent, and responsible. In this context of innovation, legal and technological communities must work together to identify, discuss, and investigate solutions to the challenges associated with these technologies.

This study aims to contribute to a technological innovation that is inclusive, transparent, and responsible [6], by surveying Judges affiliated with Portuguese Judicial Courts about the opportunities, risks, and challenges

of Digitalization, Automation, and Artificial Intelligence technologies. In this brief report, we present part of the data collected in the context of this study, focusing on *Priorities in the Digitalization and Automation of the Judicial Process; Training of Judges in the New Technological Paradigm; and Artificial Intelligence Technologies and Legal Precedent.*

2 | Methodology

The Principal Investigator (PI) of this study developed the survey on the Qualtrics platform. Subsequently, the PI tested the survey in three pilot sessions with Judges from different Portuguese Judicial Courts. The final version of the survey was approved by the Tufts University Institute Review Board (IRB) and it was distributed to Judges affiliated with Portuguese Judicial Courts via email, through a link that was available from the end of March 2023 to the end of April 2023. The results of the survey were analyzed in Python.

3 | Characterization of the Sample

A total of 279 complete surveys were recorded (\leq three incomplete responses). We consider this sample to be diverse and representative of the population of Judges in Portugal in terms of gender, career stage, and court type. There is a prevalence of female Judges, Judges in their mid and late careers, Judges currently affiliated with lower courts (Tribunais de Primeira Instância), and Judges serving in the Lisbon District (Table 1 & Figures 1 and 2).

TABLE 1: AMOSTRA DO ESTUDO

Variable	N
Gender	
Female	158
Male	117
Career	
Early (<10 years)	34
Mid (10-25 years)	114
Late (>25 years)	129
Court	
Lower Court	167
Appeal Court	78
Appeal Court of Justice	34
District	
Guimarães	37
Porto	44
Coimbra	21
Lisboa	131
Évora	46

4 | Results

4.1 Digitalization of Judicial Adjudication

Priorities in the Digitalization of Judicial Adjudication Participants in this study prioritize a judicial adjudication system that enables a better visualization of the case and the integration of a greater number of functionalities into documents (Table 2).

TABLE 2: PRIORIDADES NA DIGITALIZAÇÃO DO PROCESSO JUDICIAL

Priority	System
1	Case management system that allows for a better visualization of the case.
2	Case management system that allows for a greater number of features in documents (e.g. document editing).
3	Complete digitization of cases, including appendices.
4	Case management system where documentation is presented dynamically (e.g. hyperlinks to different documents).
5	Other

4.2 Automation of the Judicial Adjudication

Priorities in the Automation of Judicial Adjudication Participants in this study prioritize systems that are able to automatically determine court deadlines, territorial jurisdiction, and pensions (e.g. alimony) (Table 3).

TABLE 3: PRIORITIES IN THE AUTOMATION OF JUDICIAL ADJUDICATION

Priority	System
1	Determination of court deadlines
2	Determination of territorial jurisdiction
3	Calculation of pensions (e.g. alimony)
4	Determination of court costs
5	Calculation of compensation
6	Determination of the monetary value of the legal action
7	Determination of <i>legal cumulus</i>
8	Other

4.3 Technological Intervention in Justice

Judicial Adjudication that is More Favorable for Technological Intervention in Digitalization, Automation, and Artificial Intelligence Participants in this study consider that the Civil Law Adjudication is the most favorable for technological intervention, while the Criminal Law Adjudication is the least favorable to such intervention (Table 4).

TABLE 4: JUDICIAL ADJUDICATION MOST FAVORABLE TO TECHNOLOGICAL INTERVENTION IN DIGITALIZATION, AUTOMATION, AND ARTIFICIAL INTELLIGENCE

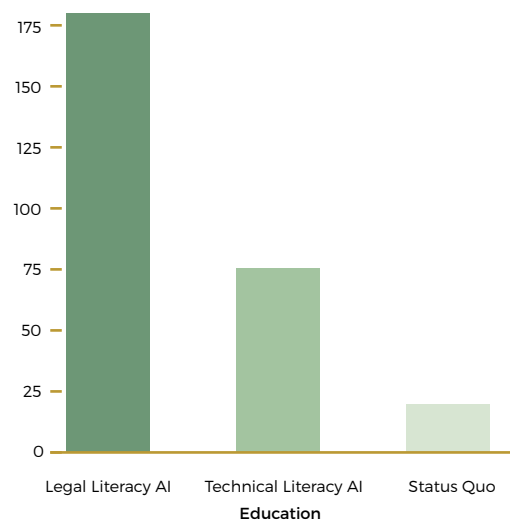
Ranking	Process
1	Civil
2	Social/Labor
3	Criminal

4.4 Training of Judges in the Novel Technological Paradigm

Training of Judges Participants in this study believe that the training of Judges in the novel technology paradigm should place a greater emphasis on improving legal literacy about AI, in order to prepare these practitioners for the legal challenges associated with these technologies. Participants also consider that this training should be offered across different phases (Law School; Judge Traineeship provided by the Center of Judicial Studies; and Continuous Education and Training, also provided by the Center of Judicial Studies) (Figure 2).

FIGURE 1: LEGAL TRAINING IN THE AI PARADIGM: TRAINING OF JUDGES

JUDGE TRAINING IN THE AI PARADIGM: LITERACY



MATERIALS TO FOSTER TECHNICAL LITERACY IN AI LITERACY

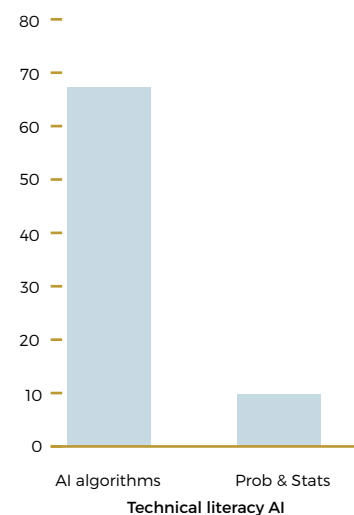
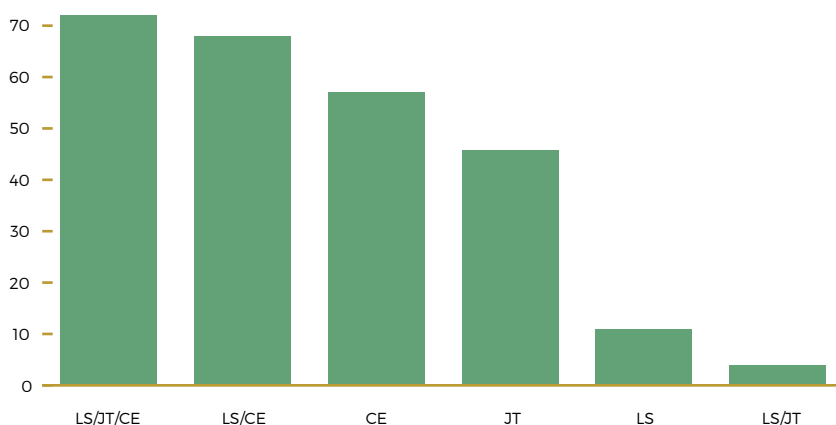


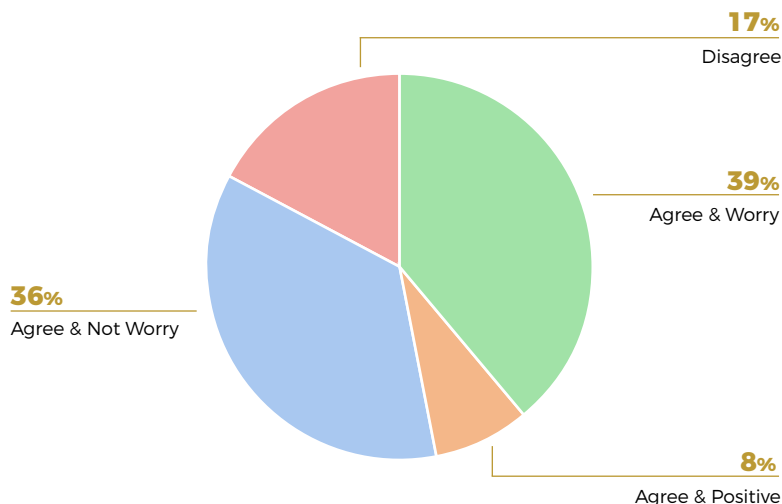
FIGURE 2: LEGAL TRAINING IN THE AI PARADIGM: TRAINING PHASE (LS STANDS FOR LAW SCHOOL; JT STANDS FOR JUDGE TRAINEESHIP; AND CE STANDS FOR CONTINUOUS EDUCATION AND TRAINING)



4.5 Judicial Precedent

Artificial Intelligence Technologies and Judicial Precedent Participants in this study believe that, considering that AI technologies identify patterns in data to formulate predictions, when applied in the legal domain, these technologies may lead to a transition to a legal model closer to Common Law, relying largely on precedent (75%) - only approximately half of these participants (39%) find this scenario concerning (Figure 3).

FIGURE 6: AI LEADING TO A TRANSITION TO A JUDICIAL MODEL GROUNDED IN JUDICIAL PRECEDENT



5 | Discussion

In this study, we investigated the perspectives of Judges currently affiliated with Portuguese Judicial Courts on the opportunities, risks, and challenges of Digitalization, Automation, and Artificial Intelligence technologies in the legal system.

On the *Priorities in Digitalization and Automation of the Judicial Process*, we consider that some systems should receive urgent attention from the entities responsible for managing technologies in the courts. We acknowledge that the implementation of some of these systems may require a high logistical and investment effort (e.g. full digitization of processes, including appendices), however, other systems seem to be relatively simple to develop and implement (e.g. correct designation of all acts mentioned in the process history, determination of court deadlines, and determination of territorial jurisdiction).

On the *Training of Judges in the New Technological Paradigm*, our results indicate that Judges believe their training should include a greater focus on improving legal literacy on AI. As an Applied Science, Law aims to provide practitioners with the tools that allow them to solve legal problems. Further research is needed to identify the best approach to incorporate these materials in the legal curricula and training, in a way that facilitates the critical thinking and exploration of the moral and ethical questions associated with AI, Data, and Algorithms [2, 3].

On *Artificial Intelligence Technologies and Judicial Precedent*, our research suggests that participants believe that AI technologies can bring significant changes to the current judicial model. These participants consider that in the future we may see an increased use of judicial precedent in the Portuguese Courts. Additional research is needed to develop the theoretical foundations of the judicial model in the novel AI paradigm.

It is important to note that this study reflects only the views of Judges, however, for technological innovation to be truly inclusive, it is necessary to listen to other legal professionals.

6 | Acknowledgements

The PI of this study makes a special acknowledgment to Judges Sandra dos Reis Luís, João Ferreira, and João Miguel Mendes, for their collaboration in the development of the survey. The PI also expresses gratitude to Judges

Sandra dos Reis Luís and João Miguel Mendes for promoting engagement in this research. The author acknowledges the Judges who participated in the testing sessions for their feedback about the survey: Ana Chambel, Cá-tia Santos, Filipe Aveiro Marques, Gabriela Cunha Rodrigues, Georgina Rodrigues, João Cura Mariano, Maria Olinda Garcia, and Pedro Soares de Albergaria. The author acknowledges the Supreme Court of Justice, represented by its President, Judge Counselor Henrique Araújo, for collaboration in this research project. Finally, the author expresses gratitude to all the Judges who completed the survey. •

REFERENCES

- [1] J. J. Baker. 2018: A legal research odyssey: Artificial intelligence as disruptor. *Law Libr. J.*, 110:5, 2018.
- [2] M. Reid. A call to arms: Why and how lawyers and law schools should embrace artificial intelligence. *U. Tol. L. Rev.*, 50:477, 2018.
- [3] F. Ryan. Rage against the machine? incorporating legal tech into legal education. *The Law Teacher*, 55(3):392–404, 2021.
- [4] F. Santoni de Sio and J. Van den Hoven. Meaningful human control over autonomous systems: A philosophical account. *Frontiers in Robotics and AI*, 5:15, 2018.
- [5] T. Sourdin. Judge v. robot: Artificial intelligence and judicial decision-making. *UNSWLJ*, 41:1114, 2018.
- [6] R. Von Schomberg. Towards responsible research and innovation in the information and communication technologies and security technologies fields. *Available at SSRN* 2436399, 2011.





COMO ESTÁ A INTELIGÊNCIA ARTIFICIAL A REVOLUCIONAR O DIREITO: UM CAMINHO SEM RETORNO



Paulo Novais

Professor Catedrático no Departamento de Informática e investigador no ALGORITMI Centre
Escola de Engenharia, Universidade do Minho

[Clique para assistir.](#)



RESUMO: A intersecção entre a Inteligência Artificial (IA) e o Direito constitui um tópico de discussão determinante nestes novos tempos. A IA está a trilhar um caminho sem retorno, com o Direito a desempenhar um papel fundamental no seu desenho, regulação e uso. Embora no passado, a IA tenha sido frequentemente associada a cenários de ficção científica, na realidade, é uma tecnologia que demonstra ser útil e que adiciona valor em muitos cenários, que vão desde as tarefas do dia-a-dia até ao trabalho, passando pelas atividades de lazer.

De facto, as “máquinas inteligentes” estão a aproximar-se do nível humano em determinadas funções fundamentais. Atualmente, a visão computacional supera a visão humana, e a tradução de idiomas encontra-se muito próxima da capacidade humana. As competências em visão, escrita (texto), imagens e fala estão praticamente ao nível humano.

A IA surge como um auxiliar na gestão da “complexidade” do mundo em que vivemos, fornecendo meios para analisar problemas e extrair informações válidas para encontrar soluções e traçar caminhos.

Existem, fundamentalmente, duas abordagens predominantes na IA. A primeira, denominada abordagem simbólica, baseia-se na representação e manipulação de símbolos e regras lógicas para a resolução de problemas, raciocínio e tomada de decisões. A segunda, conhecida como abordagem sub-simbólica ou conexionista, aprende a partir de dados, experiências e observações. Esta última é, particularmente, eficaz em tarefas que envolvem grandes volumes de dados e padrões complexos que podem ser desafiadores para expressar de uma forma simbólica. A IA oferece-nos o potencial para melhorar o modo como usamos e aprendemos conhecimento.

A IA tem desempenhado um papel decisivo no apoio ao Direito, abrangendo tarefas que vão desde a automatização de processos de inquirição, análise de contratos, até à previsão de resultados, incrementando assim a eficiência e a precisão. Adicionalmente, a IA tem, ainda,

emergido como um facilitador fundamental no acesso à justiça, proporcionando assistência jurídica automatizada e personalizada, contribuindo para a democratização do Direito ao torná-lo mais acessível ao público em geral.

Estamos a entrar numa era em que questionar uma “máquina” para obtenção de informações e conhecimento substituirá, em grande parte, a necessidade de a programar. As competências intrinsecamente humanas ganharão uma importância crescente neste cenário. A utilização de artefactos baseados em IA acarreta riscos, incluindo questões de responsabilidade, explicabilidade e transparência. A capacidade de comunicar eficazmente e de compreender o contexto será crucial para garantir que a IA interprete corretamente o contexto e forneça respostas adequadas. A emergente realidade irá potencializar os indivíduos mais proficientes, enquanto aqueles que não possuírem a capacidade de avaliar e compreender a adequação das respostas poderão não ver melhorias significativas. A importância do senso crítico, da experiência acumulada, do conhecimento adquirido e da ética na análise de respostas (soluções) serão competências valorizadas.

A Inteligência Artificial tem o potencial para expandir as competências humanas tanto ao nível individual como de comunidade de uma forma nunca vista nem sequer imaginada. Temos de ser capazes de tomar as decisões mais corretas e adequadas neste novo mundo, respeitando a centralidade do Homem e da dignidade humana. Além disso, é essencial ter em mente que a forma como as “máquinas” percecionam a realidade é de facto diferente da perceção e sensibilidade humana.

“Elas” (as máquinas) afinal não são humanas!

PALAVRAS-CHAVE: Inteligência Artificial, Aprendizagem Computacional, IA Generativa, Oportunidades e Riscos da IA, IA e o Direito.

Muito obrigado pelo amável convite, é um prazer e uma honra estar aqui no Supremo Tribunal de Justiça.

No âmbito desta comunicação, vamos abordar alguns caminhos, que consideramos que são sem retorno, que a Inteligência Artificial (IA) e o Direito estão a seguir. À medida que entramos numa era impulsionada pela IA e tecnologias digitais é essencial considerar como o Direito pode moldar e, em certa medida, regular esse desenvolvimento, definindo os seus limites e impactos.

Um dos principais problemas que a IA tem, está associado ao próprio nome e também nos muitos anos em que estivemos mais relacionados, como área científica, a uma realidade mais de *Hollywood* e de um conjunto de

cenários que foram criados pela ficção científica, que levaram a contextos (imaginários) de utilização da IA nem sempre positivos e que, de algum modo ainda estão no subconsciente das pessoas. O nome “artificial”, na medida em que a inteligência não é simples de definir, surge em contraponto ao que não é natural.

Na realidade a IA não é nada daquilo que se vê e observa nesses cenários. Obviamente, que alguma IA está a ser utilizada para fins militares, mas isso está associado a todas as formas de tecnologia, até o Direito pode ser utilizado em contextos de litigância armada.

Observemos alguns cenários. Um robô dotado de IA, porventura, desloca-se de forma mais eficiente que um humano num ambiente dinâmico como um armazém, como um hospital, por exemplo. As novas formas de interação como os assistentes pessoais na área da moda, a escolher e a adequar as nossas gostos e/ou preferências, em ambiente industrial novos sistemas que nos ajudam a garantir um elevado controle de qualidade de coloração de produtos.

Em muitos aspetos, muitos ligados aos sensores e perceção, as máquinas superam os seres humanos, o que as torna especialmente úteis em várias tarefas.

Permitam-me uma brincadeira, um jogo que faço sempre: qual dos colegas consegue identificar qual é, daquelas 4 peças, a que vai para posição onde estão os pontos de interrogação. Se demorarem mais de 2/3 segundos a responder as máquinas serão mais rápidas.

Não vos vou maçar: a resposta é o A e, porventura poderão ter identificado o porquê por diferentes formas. Uma evidente é do ponto de vista da observação visual: consegue-se identificar que falta a peça em que o meio está mais alargado. Mas sabiam que não é a única possibilidade de identificação? Se repararem com cuidado, em cada transição, o que está a acontecer é que existe um segmento de reta que se altera, que muda de posição. Por isso reparem, um problema tão simples quanto este e tem diferentes formas, diferentes perspetivas para o analisar e/ou para se obter uma solução, uma resposta. E neste? Mais uma vez perdemos, não temos capacidade de processar informação em tempo suficiente para ganhar a uma máquina neste emaranhado de movimentos...

A questão que se coloca é qual é a grande diferença entre a primeira imagem e esta segunda? O que aconteceu aqui é que a complexidade, as dimensões que estamos a utilizar são muito mais na segunda imagem.

De algum modo hoje, enfrentamos os mesmos problemas que foram analisados há séculos, mas agora analisamos com uma variedade de dimensões adicionais.

O meu avô era do Norte, agricultor e produzia vinho. As variáveis que utilizava para determinar o preço da uva eram essencialmente a produção no vale (quantidade de uva produzida nesse ano). Os meus tios, hoje, utilizam variáveis adicionais, necessitam de saber a produção que houve em outras latitudes, como o Chile, a Austrália, a Nova Zelândia, enfim num conjunto de outros países, para determinar o valor que a uva terá.

Em conclusão, a IA veio para nos ajudar a resolver o problema de ler na “complexidade”, fornecendo meios para analisar problemas e podermos retirar deles a informação válida e necessária para encontrar soluções.

Eu sigo muito uma definição recente de um dos pais desta nova Inteligência Artificial, que refere, numa visão que eu considero mais humanista da IA, que a inteligência é aquilo que, de facto, faz de nós humanos, o que nos distingue dos outros seres vivos. A IA funciona como uma extensão dessa qualidade e é nessa perspectiva que eu gostaria de continuar esta conversa.

Existem na essência duas distintas perspectivas/abordagens.

Uma abordagem simbólica que representa a verdade (o conhecimento) em áreas de conhecimento bem definidas e consolidadas. Nesses domínios, é possível transmitir conhecimento de forma estruturada, como o fazemos há séculos (recordo que existem papiros egípcios – cerca de 1500 antes de Cristo – que descrevem conhecimento médico). Essa abordagem permite programar máquinas para agir com base em regras do tipo “se isto (condições), então aquilo (ação)”.

Uma segunda abordagem muito baseada na observação de como os outros fazem é identificando semelhanças, repetições, padrões, porventura técnicas até mais adequadas de fazer uma determinada tarefa. Isso acontece, em áreas onde por exemplo os artesãos, os especialistas têm dificuldade em expressar por regras, em expressar de forma estruturada o seu conhecimento; se falarmos com um pintor ele tem dificuldade em explicar porque é que pinta (se expressa) daquela forma, mas, no entanto, qualquer um de nós olha para um quadro, por exemplo de Van Gogh, e consegue identificar que é dele.

Na essência, na abordagem simbólica podemos, se conhecemos o processo, olhar para um robô aspirador, como é o caso neste exemplo, e entender como funciona, como é que ele vira à esquerda, vira à direita e tudo isso é

passível de ser explicado porque de facto seguiu um processo bem conhecido e definido, que está a replicar. Vamos a outro exemplo, o problema do cubo mágico, na essência, partimos de um estado inicial e pretendemos chegar a um estado final conhecido. Na essência é similar ao robô que começa numa ponta da sala e tem que chegar à outra aspirando o espaço, é pelo menos isso que se espera, sabemos quais são os graus de liberdade que temos, as ações possíveis que ele pode tomar e sabemos que temos a capacidade de gerar todas as caminhos possíveis como se fosse um sistema de navegação, que temos no nosso carro, e depois por um critério qualquer, escolher qual é o caminho mais adequado. Conclusão: temos um caminho definido por um critério e podemos explicar porque é que isso aconteceu.

Há, no entanto, muitas áreas, onde não há competência humana, onde não há capacidade de explicar, onde os fenómenos mudam rapidamente. Imaginem contextos de fraude onde a criatividade humana está sempre a desenvolver novas formas, modelos baseados em grandes volumes de dados.

Um cenário exemplificativo do mercado imobiliário: temos uma base de dados com os registos de vendas de apartamentos de uma determinada zona, e por área da habitação temos o preço de venda. Usei uma regressão linear para modelar o comportamento dos dados. Modelar porque a função não tem de ser precisa, porque um comportamento não é algo de absolutamente rigoroso e exato. Podiam-me perguntar: então e aqueles dois pontos (que parecem não seguir a função modelada)? Bem, é fácil, o de cima, foi o “Cristiano Ronaldo” que comprou um apartamento em Lisboa e que naturalmente, como não é funcionário público como eu, não teve de se preocupar muito com o preço; o mais baixo é, naturalmente, o Paulo Novais que teve que negociar, regatear o preço. São *outliers*, valores atípicos que “fogem” da norma e que, como tal, o comportamento apreendido pode não modelar.

Obviamente, nem sempre os problemas são tão simples. Neste outro exemplo, representativo de um cruzamento no Brasil. Todos nós sabemos que o trânsito aumenta, porventura, das 4 da manhã até a meia-noite, é normal. Uma regressão linear para modelar seria aparentemente a solução mais adequada, mas não é porque todos nós sabemos que, pelo menos ao meio-dia há diminuição de trânsito e, provavelmente, neste caso uma regressão polinomial seria o adequado. Afinal, o que a aprendizagem automática (*Machine Learning*) faz é a procura da “função” que melhor se adequa/modela a um determinado comportamento.

Hoje, todas as nossas relações são analisadas/controladas por algoritmos. A Inteligência Artificial está, de algum modo, a influenciar as nossas

vidas. Quando acedemos ao *Google* estamos a utilizar Inteligência Artificial. Quando fazemos buscas no *Booking* para selecionar o melhor hotel estamos a usar algoritmos de IA e esses algoritmos aprendem com as nossas interações, aprendem a ajustar as nossas necessidades, as nossas preferências. Hoje, temos sistemas de Inteligência Artificial que já aprendem também pela observação, pela visão. Qualquer pessoa que jogue xadrez sabe que não é capaz de vencer uma máquina. Mas isso não é normal? As “máquinas” têm capacidades para calcular, para processar e reter, muitos superiores ao humano. Qual é o problema? O xadrez deixou de ter a beleza e o impacto que tinha? Pelo contrário: o xadrez nunca foi tão bem jogado, nunca os humanos jogaram xadrez de tão alta qualidade, porque estão a aprender também as novas estratégias que as máquinas conseguiram desenvolver por conseguirem “pensar/prever” o jogo, com muita antecipação, o que para nós é naturalmente mais difícil.

Vou, agora, focar-me um pouco mais no Direito.

Não se esqueçam que os dados são os ovos que fazem as omeletes nesta nova sociedade e sem dados não há conhecimento e sem conhecimento não há o combustível para alimentar esta nova sociedade que estamos a construir.

Nós já temos sistemas que nos permitem fazer procuras jurídicas, descoberta de provas, sistemas preditivos, a divisão de bens em caso de separação, em caso de herança (o sistema “Family winner” é um famoso algoritmo desenvolvido na Austrália, por John Zeleznikow), automatização de documentos, a simples faturação eletrónica nos escritórios dos nossos juristas, a famosa utilização do sistema COMPAS nos EUA (e também polémica), nos tribunais para avaliar a probabilidade de um arguido se tornar reincidente.

Uns exemplos daquilo que temos estado a desenvolver, por exemplo no Universidade do Minho, onde temos um mestrado de Direito e Informática: temos vindo a desenvolver sistemas para lidar com conflitos, em que aplicamos o cálculo da ZOP (zona de possível acordo) em que identificamos qual é o melhor cenário e o pior cenário de um possível, de um determinado cenário de conflito. Na essência, identificamos às partes a zona onde tipicamente podem fazer acordos, identificando jurisprudência. Mas não é isso que os colegas fazem? Ver exemplos anteriores, ver o conhecimento usado para o replicar num novo contexto similar?

Mas de repente, o mundo mudou e surgiu um “oráculo” que alguns dizem que é um terror e outros dizem que é uma revolução. Este “oráculo”, com capacidade de responder a todas as questões, parece saindo do filme *Matrix* em que *Neo* procura o oráculo quando tem dúvidas.

De que é que estamos a falar?

De um processo de aprendizagem, de grandes volumes de dados que aprende padrões (por exemplo, quando alguns de vós colocavam a fotografia e queriam que aparecesse aquela vossa fotografia pintada pelo Van Gogh ou por Dali), dos textos dos grandes escritores, pensadores, com as fontes do conhecimento humanos digitalizadas.

De algum modo, estamos a sintetizar o conhecimento humano (parte dele). Encontrando as repetições, os padrões, etc... e isso gerou o *CHATGPT*.

Eu não consigo correr e ganhar a um carro ou às motas, em particular. São muito mais rápidos que qualquer humano, mas isso faz de mim menos humano? Eu não acho. É natural que a memória de uma máquina seja maior, mais rápida e mais eficiente que a de um humano. Ela foi feita para isso, não foi feita para outra coisa. Convém, não esquecer e não confundir, estamos a falar de uma nova inteligência pouco humana mas sim máquina.

Para mim, quando falamos de inteligência não estamos a falar de quem vê melhor, de quem ouve melhor, estamos a falar de quem analisa melhor, de quem tem e usa senso comum, aquela sensibilidade que faz de nós humanos, o senso crítico, essa capacidade multidimensional de sermos humanos.

No entretanto, apareceram ferramentas para tudo e para mais alguma coisa. Ajudam-nos a escrever textos, ajudam-nos a fazer traduções. Brevemente, um dos *drones* da Guerra da Estrelas, o C3-PO, aquele mestre de cerimónias estará disponível. Isto é extraordinário! Temos, hoje, um conjunto de ferramentas que aprenderam com base naquilo que é a experiência humana, como por exemplo o *Dali-e*, essa famosa ferramenta que apareceu agora e que permite, com base na descrição de um contexto, gerar imagens. Isso não é extraordinário?

Atualmente, há um conjunto alargado de ferramentas baseadas em tecnologia de IA que devem ser utilizadas, podendo aumentar e expandir as nossas capacidades de gerar e produzir conhecimento. A justiça em Portugal já está a utilizar essas novas vias. Por exemplo, o *Guia Prático da Justiça*, ferramenta digital que permite o acesso rápido e simplificado a informação da área da justiça, é um dos primeiros exemplos de utilização, de IA, numa abordagem generativa, para responder neste caso em contextos de divórcio. E existem outros exemplos como a *Law GPT*, sistema anglo-saxónico.

Estamos a entrar numa nova era, numa era em que não vamos ter a necessitar de programar as máquinas, mas vamos ter a necessidade de interagir estas “máquinas” para obter informações/conhecimento.

As competências que fazem de nós humanos vão ser cada vez mais im-

portantes neste cenário, porque só quem for capaz de interrogar estas máquinas de uma forma adequada é que vai ser capaz de tirar proveito delas. Porventura, vamos entrar numa sociedade ainda mais desigual porque o que acontece é que serão os especialistas, neste caso do Direito, que serão capazes de utilizar essas ferramentas de forma adequada e não o cidadão comum, porque, obviamente, eu não sou capaz, como vou mostrar a seguir, de detetar erros, porque me falta esse senso crítico.

Explorando o *CHATGPT* lembrei-me de lhe fazer algumas perguntas. Por exemplo onde é que faz sentido utilizar o modelo *GPT*?

E, de facto, ele identifica coisas interessantes. Continuo a entender que nos ajuda na esfera jurídica, mas há uma outra dimensão, pode analisar contratos (verificando se foram identificados incumprimentos, fazer análise preditiva, entre outras).

Usei um artigo de um colega americano, solicitei a sua tradução para português e que ele destacasse as principais ideias do texto. Funcionou e bem!

De seguida, solicitei a geração de uma ação, explicando que tive um acidente de viação na minha cidade, Guimarães, e reparem ele foi capaz de gerar um texto em que, de algum modo, explica um contexto. É rigoroso, é adequado? Porventura não, mas é um ponto de partida, não começamos do zero.

Lembrei-me de fazer uma brincadeira e pedir para ele me redigir o meu testamento. Neste caso confesso que deixava tudo à minha mulher. Reparar que ele (o *CHATGPT*) até identificou que ela é amada por mim, muito embora eu não lhe tenha dito no texto, mas naturalmente se se deixa algo a alguém, é porque se nutrem sentimentos. A tal correlação nos textos que normalmente estão associados em testamentos surge, aqui, refletida. Está correto, está errado? Bem, como ponto de partida, mais uma vez, parece-me certo. Claro que estes são exemplos triviais.

Há riscos? Há muitos riscos na utilização dos artefactos baseados em IA, obviamente que não vamos negar isso.

A questão da responsabilidade: será que os conselhos que umas máquinas nos fornecem são ou não adequados e fidedignos? Bem, eu não sei se sabem os riscos que vocês têm quando utilizam, por exemplo, uma ferramenta como um Antivírus. Algum de vós foi ver as contraindicações de utilização desse tipo de *software*? Se o tivessem feito nenhum de vós o utilizava, porque os “criadores” não se responsabilizam por nenhum erro.

No cenário de um veículo autónomo, imaginem que ele (o veículo) pode

aprender no decurso de uma viagem, tal e qual como o *CHATGPT* aprende com as nossas perguntas, e aprende algo que pode colocar as nossas vidas (passageiros do veículo) em perigo. Aceitariam viajar num veículo destes? Eu não!

A transparência é um dos problemas, porque estamos a falar em modelos estatísticos. Será que aceitam uma decisão baseada numa explicação do tipo “porque 90%”?

Os vírus do futuro serão fruto da injeção de dados maliciosos, deixa de ser no processo, afetando os dados que usamos para suportar as decisões. Quero aqui alertar que nem sempre nós (humanos) sabemos explicar as nossas decisões, muitas vezes só a *posteriori*, depois de ter decidido, é que eu tento encontrar/procurar uma explicação racional para a minha decisão.

Acima de tudo temos de ter consciência que, para esta nova IA, os conhecimentos de linguagem, de comunicação, serão fundamentais, porque a habilidade de comunicar, perceber e saber expor as questões que temos serão basilares para que a IA, a máquina, interprete corretamente e adequadamente o contexto e assim possa dar respostas mais adequadas. Mais uma vez, os ovos e a sua qualidade serão fundamentais para se fazerem omeletes.

Outra questão que se coloca é a saber quem vai poder tirar proveito destes novos artefactos, destes novos ventos.

Na verdade, quem poderá aproveitar estes instrumentos é quem tiver a capacidade de avaliar a qualidade das respostas. Esta nova realidade tornará os bons muito melhores, mas os que não têm essa capacidade não vão melhorar, porque obviamente vão cometer erros, porque não vão ser capazes de avaliar a adequação das respostas. Tal e qual como eu no caso do contrato ou do testamento que illustrei anteriormente.

É essencial ter senso crítico para analisar respostas sob diferentes perspetivas e avaliar se incluem todos os aspetos relevantes. A experiência e o conhecimento desempenham um papel fundamental nessa capacidade, assim como a ética. Uma utilização ética e responsável é crucial. Como engenheiros, podemos incorporar responsabilidade ética no desenvolvimento desses artefactos para evitar situações problemáticas, como no caso do *CHATGPT*, onde respostas ofensivas são evitadas devido a uma programação ética.

Uma mensagem de otimismo: eu estou convencidíssimo que esta é uma revolução que veio para ficar, é um caminho sem retorno nas diferentes áreas de atuação da humanidade e que vai expandir, de facto, as nossas

capacidades como pessoas, como indivíduos no nosso dia-a-dia, nas atividades de lazer e no nosso trabalho de uma forma absolutamente disruptiva e, também, creio que como comunidade, como sociedade, vai expandir as nossas capacidades como espécie.

Cabe-nos a nós tomar as decisões mais corretas e adequadas no mais elementar respeito pela dignidade humana e nunca esquecer que de um lado temos a realidade em que vivemos e, do outro lado, temos a forma como as “máquinas” veem essa realidade, que são necessariamente diferentes. “Elas” (as máquinas) afinal não são humanas! •

Muito obrigado!



APLICAÇÕES DE GRANDES MODELOS DE LINGUAGEM NA ÁREA JURÍDICA



Arlindo Oliveira
INESC-ID / Instituto Superior Técnico

Clique para assistir.



1 | Introdução

A integração de tecnologias digitais têm-se mostrado essencial para otimizar processos, aumentar a eficiência e proporcionar uma administração mais justa e eficaz da justiça. A utilização de tecnologias digitais como bases de dados e processadores de texto é já uma prática corrente, mas não extrai ainda todo o valor possível das tecnologias digitais actuais, nomeadamente as de inteligência artificial que se desenvolveram nos últimos anos.

Nesse contexto, os grandes modelos de linguagem, como o *ChatGPT* ou o *GPT-4*, têm emergido como ferramentas poderosas capazes de transformar a prática jurídica e tornar mais eficientes a elaboração, análise e sumarização de documentos. Neste curto texto, as tecnologias usadas são descritas de uma forma sintética e são identificadas possíveis aplicações desses modelos, em áreas que incluem a análise e indexação de jurisprudência, a pesquisa de legislação, a comparação de decisões judiciais, a sumarização de processos e a identificação de irregularidades em documentos legais. São ainda elencadas aplicações mais específicas em área operacionais relacionadas com os processo judiciais.

2 | Grandes Modelos de Linguagem

Um modelo de linguagem é uma aplicação computacional que tem a capacidade de entender e gerar texto de maneira coerente e contextualmente

relevante. Este tipo de modelo é geralmente baseado em uma arquitetura tecnológica conhecida como “transformador”. Os transformadores são uma classe de modelos de redes neurais, uma tecnologia específica de aprendizagem automática, que foram projectados para processar sequências de símbolos, como as palavras de uma frase ou parágrafo. Os transformadores incluem camadas de auto-atenção, que permitem que o modelo pondere a importância de diferentes partes da sequência ao realizar tarefas de processamento de linguagem natural. A grande vantagem dos transformadores é a sua capacidade para capturar relações complexas em contextos linguísticos extensos. Os modernos modelos de linguagem são parametrizados com centenas de milhares de milhões de parâmetros, que são derivados durante o processo de treino.

A metodologia de treino destes modelos geralmente envolve grandes quantidades de dados textuais, variados e obtidos de diversas fontes que incluem livros, artigos, *websites* e outros textos disponíveis na *internet*. O GPT-3, um dos primeiros modelos de grande dimensão, disponibilizado em 2020, foi treinado com dados textuais que exigiriam a um ser humano 5000 anos de leitura contínua, 24 horas por dia. A grande dimensão do conjunto de dados usado para treinar os modelos justifica, em parte, a capacidade destes modelos para processarem e entenderem língua natural, em diversas línguas e idiomas. Os símbolos são “embebidos” num espaço vectorial, com algumas centenas de dimensões (tipicamente 1024 dimensões) o que permite fazer extrapolações entre línguas, tempos verbais, género, culturas, localização geográfica e outras dimensões relevantes, que são identificadas automaticamente pelo processo de “*embedding*”.

Durante o treino, o modelo ajusta os seus parâmetros de acordo com os padrões linguísticos presentes nos dados, por forma a melhorar a sua capacidade de prever o próximo símbolo. A capacidade de prever o próximo símbolo está, de facto, intimamente relacionada com a capacidade para explorar modelos internos de linguagem e do mundo, o que permite aos modelos fazer extrapolações, inferências e relações que extravasam, em muito, os dados com os quais foram treinados. Os grandes modelos de linguagem têm um potencial significativo em diversas aplicações, que vão desde assistentes virtuais e tradução automática até à geração de conteúdos. A sua capacidade de compreender contextos complexos e aprender a partir de grandes volumes de dados textuais permite que esses modelos gerem texto coerente e informativo numa grande variedade de situações.

No entanto, é importante destacar algumas limitações. Esses modelos

podem ocasionalmente gerar respostas incorrectas ou enviesadas, refletindo os padrões presentes nos dados de treino mas uma limitada compreensão das questões específicas. Além disso, a compreensão desses modelos é relativamente superficial, uma vez que eles não possuem uma compreensão da semântica dos textos, dependendo principalmente de padrões estatísticos aprendidos durante o treino. É importante manter presente que os modelos actualmente existentes são puramente modelos estatísticos, que geram a próxima palavra de acordo com a ocorrência das palavras anteriores e que, portanto, não dão respostas que sejam, garantidamente, factualmente correctas. É no entanto de esperar que os próximos modelos venham a ser mais capazes e menos sujeitos ao fenómeno da “alucinação”, que se refere à invenção de factos, plausíveis mas inexistentes, que afecta ainda fortemente muito dos actuais modelos de linguagem.

3 | Áreas de Aplicação

As potenciais áreas de aplicação dos grandes modelos de linguagem na área jurídica são numerosas e multifacetadas e incluem, entre outras:

- **Análise e Indexação de Jurisprudência:** os grandes modelos de linguagem podem ser empregados para analisar jurisprudência de forma rápida e eficiente. Ao utilizar algoritmos de processamento de linguagem natural, esses modelos conseguem identificar padrões, tendências e relações em decisões judiciais, facilitando a categorização e indexação de casos. Isso não apenas acelera a pesquisa de informação relevante, mas também contribui para uma compreensão mais abrangente das interpretações legais ao longo do tempo.
- **Pesquisa de Legislação Relevante:** a busca por legislação relevante é uma tarefa complexa e muitas vezes demorada. Os grandes modelos de linguagem podem ser empregados para criar motores de busca jurídicos mais sofisticados, compreendendo nuances semânticas e contextuais. Esses motores podem responder a consultas jurídicas de maneira mais inteligente, fornecendo resultados mais precisos e economizando tempo dos juristas na pesquisa de normas e leis específicas.
- **Comparação de Decisões Judiciais:** ao utilizar técnicas avançadas de análise de texto, os modelos de linguagem podem realizar comparações detalhadas entre decisões judiciais. Isso é particularmente útil em casos similares, pois ajuda a identificar discrepâncias e similaridades, permi-

tindo uma fundamentação mais consistente e uma compreensão mais profunda da jurisprudência vigente.

- **Sumarização de Processos e Documentação Associada:** a extensão e o elevado número de documentos legais associados a processos muitas vezes sobrecarrega os profissionais e dilata a duração dos processos. Os modelos de linguagem podem ser treinados para sumarizar processos judiciais e documentos associados, destacando informações cruciais e eliminando detalhes redundantes. Isso simplifica a compreensão do caso, permitindo uma análise mais eficiente e acelerando o processo decisório.
- **Identificação de Irregularidades em Documentos Legais:** a detecção de irregularidades em documentos legais é uma área crítica no campo jurídico. Os modelos de linguagem podem ser configurados para analisar contratos, decisões judiciais e outros documentos em busca de inconsistências, omissões ou cláusulas problemáticas. Essa capacidade de identificação precoce de potenciais problemas contribui para a mitigação de riscos legais e para a tomada de decisões mais informadas.

Além das aplicações mencionadas anteriormente, os grandes modelos de linguagem têm o potencial de desempenhar um papel significativo no apoio operacional aos tribunais e à investigação criminal, oferecendo ferramentas inovadoras e potenciadoras de uma maior eficiência.

Algumas outras aplicações relevantes neste aspecto incluem:

- **Elaboração de Documentos Jurídicos:** os grandes modelos de linguagem podem ser usados para auxiliar na redação de documentos jurídicos, como petições, pareceres legais e contratos. Ao oferecer sugestões de linguagem, garantindo a consistência e a conformidade com normas legais, esses modelos podem agilizar o processo de redação e melhorar a qualidade dos documentos produzidos.
- **Previsão de Resultados Judiciais:** através da análise de padrões presentes em decisões judiciais passadas, os modelos de linguagem podem gerar previsões sobre possíveis resultados em casos similares. Embora a previsão de resultados judiciais seja complexa, esses modelos podem fornecer uma avaliação inicial, ajudando advogados e partes envolvidas a tomar decisões informadas sobre estratégias legais.
- **Análise de Provas e Documentos:** na investigação criminal, a análise de grandes volumes de provas e documentos é uma tarefa difícil, pela sua complexidade e dimensão. Os grandes modelos de linguagem podem ser

empregues para identificar padrões, relações e informações relevantes em documentos, facilitando a investigação e acelerando a análise de casos complexos.

- **Tradução Jurídica:** em casos que envolvem múltiplos idiomas, os modelos de linguagem podem facilitar a tradução jurídica, garantindo uma compreensão precisa e consistente de documentos legais. Isso é crucial em investigações internacionais e em situações em que a colaboração entre jurisdições é necessária.
- **Análise de Sentimento em Textos Jurídicos:** a compreensão dos sentimentos expressos em textos jurídicos pode ser valiosa na antecipação de futuros resultados. Os modelos de linguagem podem identificar sentimentos, polaridades e nuances nas declarações e argumentos apresentados, auxiliando na preparação estratégica.
- **Triagem de Casos e Priorização de Investigação:** ao analisar automaticamente informações relacionadas com casos, os modelos de linguagem podem contribuir para a triagem e priorização eficiente de investigações criminais. Isso é especialmente útil em situações onde os recursos são limitados e é crucial focar nos casos mais relevantes e urgentes.
- **Educação Jurídica e Formação:** os modelos de linguagem podem ser utilizados na criação de materiais educativos interativos, simuladores de casos e programas de formação para profissionais do direito. Essas ferramentas podem melhorar a formação prática, oferecendo experiências de aprendizado mais dinâmicas e realistas.

4 | **Sumário e Conclusão**

O potencial para a aplicação dos grandes modelos de linguagem na área jurídica é vasta e diversificada. Ao apoiar tribunais e investigações criminais, essas tecnologias têm o potencial de transformar fundamentalmente a prática jurídica, tornando-a mais eficiente, acessível e informada. No entanto, é crucial que essas ferramentas sejam implementadas de maneira ética, transparente e em conformidade com os princípios legais, garantindo a confiabilidade e a integridade do sistema jurídico.

A integração de grandes modelos de linguagem na área jurídica tem o potencial para revolucionar a forma como os profissionais do direito lidam com informações e processos. As diversas áreas de aplicação elencadas no presente texto apontam possíveis direções para futuros desenvolvimentos que são já exequíveis, de um ponto de vista estritamente técnico. Não exis-

tindo obstáculos de natureza técnica, o principal desafio futuro reside na integração eficaz, ética e responsável dessas ferramentas, garantindo que a inovação tecnológica se alie à preservação dos valores fundamentais do sistema judicial. A questão de integração com os sistemas existentes é particularmente crítica, uma vez que poderá enfrentar desafios de natureza técnica, cultural e legislativa, que limitarão seguramente a velocidade a que a adoção terá lugar. •



JUIZ IA: FUTURO, UTOPIA OU DISTOPIA?



Vera Lúcia Raposo
Professora Auxiliar, NOVA School of Law

Clique para assistir.



PALAVRAS-CHAVE: Inteligência artificial; Algoritmo; Juiz; Tribunal; Justiça.

Muito obrigada, boa tarde a todos!

Sei que não é fácil estarmos aqui a esta hora e por isso mesmo forcei-me a pensar numa questão que fosse suficientemente atractiva e, eventualmente, provocatória, para vos manter a todos atentos. Eis a questão que espero debater convosco aqui: Será que os Juizes, muitos daqueles que estão na minha frente, irão algum dia, num futuro mais ou menos próximo, ser substituídos pela IA?

Sei que algumas das questões que vou lançar para a discussão são ligeiramente provocatórias, mas tenham paciência comigo. Há uma finalidade no meio de tudo isto.

Há uns meses atrás, um juiz da Colômbia tornou-se conhecido porque teve a audácia de dizer em público aquilo que eu desconfio que se passe com frequência em privado: confessou que tinha utilizado o *ChatGPT* para redigir uma sentença. Com esta revelação escancarou-se a Caixa de Pandora. Digo ‘escancarou-se’ porque já estava aberta. A questão já estava em cima da mesa, agora tornou-se inevitável falar dela.

Por conseguinte: a pergunta que vos trago aqui é uma que me tem ocupado nos últimos meses: Será que a IA vai substituir os humanos nos mais diversos níveis? A mim, como professora de direito? Aos médicos e engenheiros e, obviamente, aos magistrados que aqui estão?

Vou trazer-vos à discussão uma série de argumentos que têm sido avançados para afirmar que isto nunca irá acontecer, ou seja, que é absolutamente impossível que a IA venha a substituir um juiz. A minha proposta é que em conjunto os analisemos e, eventualmente, os refutemos. Por força da vossa função decisória os senhores sabem melhor do que eu que às vezes as aparências iludem e é preciso desmontar os argumentos. É exactamente isso que vos peço: desmontemos estes argumentos.

O primeiro grupo de argumentos que vos trago tem a ver com as características intrínsecas da IA, as quais parecem demonstrar, pelo menos numa primeira abordagem, que, de facto, a substituição é qualquer coisa impossível de acontecer. Passemos a analisá-los.

No âmbito desta lista de argumentos, a primeira grande objecção prende-se com as próprias dificuldades no treino de um sistema de IA que seja suficientemente bom para poder desempenhar uma função de tanta complexidade e de tanta responsabilidade como é a de um juiz. A questão coloca-se pela seguinte razão: conseguir treinar um bom sistema de IA é tarefa complicada. A IA precisa de quantidades massivas de dados, tal como nós precisamos de comida e de bebida. Sem comida e, mais do que isso, sem bebida, perecemos. Da mesma forma, se os dados não forem bons ou não existirem em quantidade suficiente, não é de admirar que o resultado não venha a ser bom. Na gíria isto é conhecido como ‘garbage in, garbage out’.

Dou-vos o seguinte exemplo para terem uma ideia mais concreta: para um sistema de IA conseguir reconhecer com 95% de precisão a imagem de um gato, é preciso que o sistema primeiro visualize 100 mil imagens de gatos. É uma quantidade imensa, sobretudo considerando que estou a falar de uma coisa simples, ou seja, reconhecer um gato. Imaginem agora quão mais complexo se torna – e a quantidade brutal de dados (petabytes de dados) que são necessários – para conseguir decidir questões complexas como aquelas que os senhores decidem todos os dias.

Sucedem que dados não faltam. Dados é o que mais há por esse mundo fora. Vivemos numa sociedade digital e, por conseguinte, estamos absolutamente rodeados por dados.

Não é um problema de falta de dados. É um problema de quadro legal para usar esses dados. Como sabem, na União Europeia temos um ‘pequeno’ obstáculo ao tratamento de dados pessoais: o Regulamento Geral da Protecção de Dados (RGPD). É considerada a mais exigente lei de protecção de dados pessoais em todo o mundo. Supostamente, o *standard of care* nesta matéria. A mais clara consequência do nível de exigência que impõe são os

muitos entraves que coloca em tudo o que diga respeito ao tratamento de dados pessoais, sendo que, obviamente, o uso de dados para treino de um sistema de IA é uma forma de processamento de dados, sendo muitos desses dados considerados dados pessoais.

Para tentar contornar esta questão, várias soluções têm sido apontadas.

A primeira delas é deixar de utilizar dados pessoais, dado que só a esses que se aplica o RGPD. Em alternativa, poder-se-ia usar dados não pessoais ou dados anonimizados para treinar a IA. Estes dois conceitos podem parecer a mesma coisa, mas são coisas diferentes. Um dado não pessoal é aquele que nunca, pela sua natureza, teve qualquer ligação a uma pessoa humana: dados sobre o clima, sobre a intensidade do trânsito numa estrada, sobre a temperatura da água nos oceanos. Estes são exemplos de dados não pessoais porque (pelo menos aparentemente) não os conseguimos reportar a uma pessoa humana identificada ou identificável. Ao invés, um dado anonimizado é um dado que começou por ser um dado pessoal, ou seja, referente a uma pessoa humana, mas que por via de técnicas de anonimização se tornou anonimizado. Qualquer deles cai fora do âmbito de aplicação do RGPD, pelo que recorrer a estes dados – e não a dados pessoais – parece uma boa forma de contornar o problema.

Contudo, não é. Passo a explicar.

Em primeiro lugar, não acho que estes dados caiam assim tão facilmente fora do âmbito de aplicação do RGPD. O problema nasce pela própria definição de dado pessoal que encontramos no artigo 4.º, n.º 1, do RGPD. O artigo 4.º contém as definições de conceitos básicos para o RGPD e no número 1 encontramos a definição de ‘dado pessoal’: ‘Informação relativa a uma pessoa singular identificada [e agora o problema] ou identificável’. Como refere Nadezhda Purtova, num dos artigos que eu pessoalmente acho um dos melhores artigos sobre o RGPD, ‘The Law of Everything’, hoje em dia, praticamente todos os dados, por mais distantes que pareçam de uma pessoa humana, acabam por ser dados pessoais. Utilizando as técnicas de IA e tendo uma boa quantidade de informação adicional para fazer as correlações e descobrir padrões, praticamente qualquer dado, até o tal dado sobre os oceanos ou o clima, consegue ser ligado a uma pessoa humana. A conclusão desta autora é que hoje praticamente não existem dados não pessoais e eu estou em crer que para aí caminhamos.

Em segundo lugar, surgem questões relacionadas com os dados anonimizados. Há uns anos atrás, as empresas anonimizavam os dados com um sentimento de segurança, acreditado que o RGPD deixava de se lhe aplicar. Hoje em dia sabemos que a anonimização absoluta é utópica, porque

utilizando técnicas de IA, por meio de uma espécie de engenharia inversa – *reverse engineering* – conseguimos reestabelecer aquela ligação que o dado inicialmente tinha com a pessoa. Isto foi, aliás, já reconhecido pelo Comité Europeu da Proteção de Dados, aconselhando os responsáveis pelo tratamento dos dados em certos domínios específicos, tal como dados de saúde e dados biométricos, a seguir os trâmites do RGPD, mesmo que os dados estejam (aparentemente) anonimizados. A título de exemplo, recorro ao incidente ocorrido nos EUA, quando os jornalistas quiseram consultar os dados fiscais do Presidente Trump que, obviamente, estavam superanonimizados, sendo que, porém, os jornalistas conseguiram desanonimizá-los através da IA.

Em terceiro lugar há questões que se prendem com a utilidade dos dados anonimizados. Cito um autor – Paul Ohm – que, há mais de uma década, disse uma verdade que me parece ainda hoje inabalável: os dados ou são perfeitamente anónimos ou são úteis. Não podem é ser as duas coisas.

Por conseguinte, usar dados não pessoais ou dados anonimizados não me parece a melhor via para ultrapassar os limites do RGPD. Porém, não é este o fim da história. Há uns anos apareceu outra possibilidade, aparentemente ‘milagrosa’: os dados sintéticos. Há quem diga que são dados falsos. Não são. São dados eles próprios gerados por IA com base em dados pessoais. A ideia inicial era que os dados sintéticos não eram dados pessoais e, por conseguinte, estariam fora do RGPD. Se assim fosse, poderíamos desenvolver sistemas de IA com grande margem de liberdade. Porém, hoje também já sabemos que assim não é. Das conversas que tenho com os *expert* de IT resultam dois tipos de opiniões: 1) Os que afirmam convictamente que os dados sintéticos são ainda dados pessoais; 2) Os que utilizando a dupla negativa, não podem garantir que não sejam dados pessoais. Em qualquer dos casos, mesmo nesta segunda hipótese, a via dos dados sintéticos parece-me pouco credível. Um outro problema prende-se com a utilidade destes dados, dado que com os dados sintéticos muito pouco se pode fazer em termos de treino de IA, de modo que não é uma solução viável.

Nos últimos tempos, porém, surgiu uma possível via. É que finalmente a Comissão Europeia percebeu que o RGPD funciona bem em teoria, mas na prática, em certos domínios, causa mais problemas do que resolve, nomeadamente pelos enormes entraves que coloca à utilização, e sobretudo à reutilização, dos dados. Ou seja, não é uma questão de escassez de dados, mas de restrições à sua utilização. Contudo, em vez de alterar o RGPD, coisa que não parece que vá acontecer num futuro próximo, decidiu-se pela criação de espaços de dados, as chamadas ‘bolhas de dados’. Basicamente, estamos a falar de um espaço circunscrito – as tais ‘bolhas de dados’ (*data*

spaces) – onde os dados podem ser mais facilmente usados, partilhados, reutilizados, em suma, processados. Estes *data spaces* alegadamente vão nascer em vários domínios estratégicos (nunca se sabe se a justiça não será um deles) e o primeiro já está neste momento em vias de ser aprovado. Trata-se do Espaço Europeu de Proteção de Dados, que tem como principal objectivo facilitar o uso de dados pessoais no treino de sistemas de IA. Se continuarmos por esta via, possivelmente teremos aqui maior flexibilidade e, por conseguinte, conseguiremos mais facilmente treinar bons sistemas de IA, sistemas suficientemente credíveis para poderem eventualmente no futuro conseguir substituir um juiz.

Segundo grande argumento contra a substituição do juiz pela IA: os vieses.

O viés é uma ideia preconcebida, isto é, um preconceito que nos leva a actuar de uma determinada forma, eventualmente conduzindo no pior dos cenários a uma decisão discriminatória.

A tecnologia não tem viés, é absolutamente neutra. Porém, a tecnologia é feita por humanos e, por conseguinte, nós, os humanos, transmitimos-lhe os nossos vieses. A transmissão dos vieses processa-se por via da codificação: quem escreve os códigos são os humanos que, por conseguinte, inadvertidamente passam as suas próprias ideias pré-concebidas para o código, fazendo com que o algoritmo actue de determinada forma.

A outra grande causa dos vieses são os dados. Muitos dos senhores terão filhos e, se bem se recordam, ao educar uma criança a sua mente é uma tela em branco. Se ensinarmos às crianças que as pessoas com uma determinada cor de pele, com uma determinada religião ou uma determinada orientação sexual, se comportam de certa forma, a criança crescerá a pensar assim e tornar-se-á num adulto que decide com base nessas ideias pré-concebidas. É o mesmo com a IA. Se eu alimentar a IA com dados que a levam a deduzir que pessoas com determinadas características se comportam de uma determinada forma, as decisões da IA vão ser nesse sentido, não esperem outra coisa porque foi isso que a IA aprendeu.

Dou-vos um exemplo concreto: nos Estados Unidos usa-se um sistema de IA preditiva chamado COMPAS para calcular a probabilidade de reincidência. É muito utilizado, por exemplo, para tomar decisões sobre prisões preventivas, libertações antecipadas, ou sobre a medida da pena. O sistema foi treinado com os registos criminais da população prisional norte-americana e com as respostas que os reclusos deram a um questionário que lhes foi distribuído. Primeiro problema: a população prisional norte-americana, como bem sabem e está descrito em inúmeros estudos, é maioritariamente afro-americana ou de outras minorias. Existe uma desproporção brutal en-

tre o que é a população prisional e o que é a população total norte-americana, sendo que os afro-americanos e outras minorias ocupam boa parte dos lugares nas prisões. A IA – que é alimentada com estes registos – aprendeu que as pessoas com estas características são mais propensas a cometer crimes. Eis aqui o viés! Segundo problema: o dito questionário tem perguntas do tipo ‘é errado roubar para comer?’. Ora bem, a resposta a esta questão depende muito do meio socioeconómico de onde cada um de nós vem, nomeadamente se alguma vez na nossa vida passámos fome ou não. Mais uma vez, demonstram os estudos que boa parte da população de estratos económicos mais baixos nos Estados Unidos é composta por afro-americanos e outras minorias, pelo que estas pessoas deram reiteradamente uma resposta diferente à pergunta. Eis aqui, de novo, o viés! O resultado é que o COMPAS sistematicamente aponta para taxas de reincidência mais elevadas para afro-americanos do que para caucasianos.

Se eu parasse aqui a minha argumentação, todos concordaríamos que este é um argumento de peso para não utilizarmos a IA como substituto do juiz: a IA tem vieses e conduz a resultados discriminatórios. O juiz não.

Porém, analisemos a questão com mais detalhe utilizando de novo o COMPAS como *case study*. Se é verdade que boa parte da população prisional pertence a estas minorias, não foi a IA que meteu estas pessoas na prisão (tanto mais que o uso da IA é bem mais recente), foram os juízes humanos, predominantemente caucasianos, que tomaram estas decisões. A partir daqui só há duas conclusões: ou bem que adoptamos uma espécie de revivalismo das teorias de Lombroso e achamos que as pessoas de determinado aspecto/grupo étnico são mais propensas a cometer crimes, o que obviamente ninguém defenderá; ou então temos que reconhecer que, de facto, os juízes humanos têm vieses, não conscientemente, mas inadvertidamente, pela simples razão de que tal faz parte da nossa natureza. Estes vieses conduzem a decisões discriminatórias.

Nem sequer é preciso analisar casos de discriminação para identificar vieses. Recentemente li um artigo onde o autor dava alguns exemplos de vários elementos aparentemente irrelevantes que condicionam a decisão do juiz (na verdade, não só de juízes, mas de todos nós, humanos). Refiro-me a elementos como os nossos valores pessoais e as nossas experiências de vida. Podemos não nos aperceber, mas são factores que condicionam a forma como pensamos nos problemas. Outros elementos são um pouco mais curiosos, tais como a hora do dia (parece que tomamos decisões diferentes consoante a hora do dia, por exemplo se for a hora do almoço e não tivermos ainda comido a decisão é diferente do que seria se tomada noutra altura)

e a fadiga na decisão (se já tivermos tomado muitas decisões decidimos de forma diferente). Deixei para o final um factor que eu acho absolutamente delicioso: a atractividade dos indivíduos envolvidos. O que estou a dizer não é que o juiz absolve ou condena o arguido porque o acha mais ou menos bonito; o que digo é que somos humanos e demonstram os estudos que o nosso cérebro humano age de uma forma que não conseguimos controlar, reagindo de forma diferente consoante o grau de atractividade que nos desperta a nossa contraparte. Isto é um facto científico. Não conseguimos controlar este ‘condicionamento’ porque também nem conseguimos identificar quando é que estamos a ser controlados por eles.

Daqui se conclui que nós, os humanos, também temos os nossos vieses, os quais nunca vamos conseguir apagar... sob pena de deixarmos de ser humanos.

Será que conseguimos resolver a questão dos vieses da IA?

Não podemos resolver o problema na totalidade, mas conseguimos atenuá-lo com melhor codificação e alimentando a IA com mais e melhores dados. Contudo, nunca iremos ter uma IA completamente livre de vieses, assim como nunca vamos conseguir juiz IA’, prende-se humanos completamente livres de vieses.

O terceiro grande argumento, que parece servir de motivo cabal para não substituir o juiz humano por um ‘juiz IA’, prende-se com o risco de decisões erróneas. De facto, os vieses não são a única causa de erros na IA. Outro problema prende-se com a sua ‘natural estupidez’.

Começo por notar que a IA tem uma performance melhor do que os humanos em várias áreas: consegue armazenar informação como nós não conseguimos e consegue identificar padrões como nós não conseguimos. Porém, no fundo é incrivelmente estúpida, ou seja, a IA consegue ser facilmente enganada. Por exemplo, no reconhecimento de imagens está comprovado que, por vezes, confunde ondulações de tom azul, amarelo e laranja com uma estrela do mar; assim como também no reconhecimento de imagens muitas vezes confunde um gato com um sapo.

Estes são exemplos reais, mas anedóticos. Deles não vem mal ao mundo, porém, estes mesmos erros podem ser replicados – e com bem maior gravidade – nas questões verdadeiramente importantes da vida, ou seja, em questões que têm impacto decisivo na vida das pessoas que têm casos em tribunal. Se pensarmos em quão dramáticos poderão ser alguns dos erros cometidos por um suposto ‘juiz IA’ no âmbito judicial, teremos que concluir que nunca poderá a IA assumir o papel do juiz.

Contudo, não nos iludamos. Se é certo que um humano consegue dife-

reconhecer um gato de um sapo, também é certo que os humanos conseguem cometer erros básicos nas suas decisões. Todos os humanos o fazem e certamente que os juizes também. Somos humanos, não conseguimos evitar os erros básicos, assim como a IA, pelo menos neste momento, não os consegue evitar.

Um outro leque de argumentos que parece militar contra a substituição do juiz humano pela IA prende-se com aquelas especificidades que são profundamente humanas.

Por exemplo, a capacidade de ‘pensar fora da caixa’. Um juiz pode conhecer todas as teorias sobre o bem jurídico, conhecer os códigos de trás para a frente, dominar toda a jurisprudência e toda a doutrina e, no entanto, haverá sempre casos que o juiz só consegue resolver se pensar fora da caixa. Já não vivemos no mundo do silogismo jurídico e ainda bem. Pensar fora da caixa é aquele rasgo de criatividade que em Portugal chamamos de ‘desenrasca’.

A IA não consegue fazer isto porque a IA é, apesar de toda a sua complexidade, um algoritmo e como algoritmo que é funciona sempre segundo a lógica ‘se-então’. É claro que os sistemas de IA que nós temos hoje em dia não são só algoritmos assim tão simples. São, pelo contrário, sistemas incrivelmente complexos, formados por teias complexas e densas de algoritmos com milhares de milhões de possibilidades. Porém, em última instância a IA funciona sempre com base nas regras. Nela não há lugar para o imprevisto, porque o imprevisto – que exige pensar fora da caixa – não consegue entrar na lógica do algoritmo. Por conseguinte, este é um ponto para os humanos.

O outro ponto a favor dos humanos é a intuição. Durante muito tempo pensou-se que o direito era profundamente lógico e racional, quase como se fosse uma variante da matemática, pelo que seria um domínio onde a intuição não teria lugar. Porém, hoje em dia, os estudos comprovam o papel da intuição na tomada de decisões judiciais. Não só comprovam, como reconhecem o valor que a intuição tem no julgamento judicial. Ora, a intuição é uma característica puramente humana, a máquina não tem intuição. Logo, outro ponto para os humanos.

Outra característica puramente humana é a empatia, ou seja, a inteligência emocional. A empatia é absolutamente indispensável cada vez que lidamos com outros humanos, em qualquer profissão, em qualquer situação da nossa vida. No contexto de uma audiência judicial a empatia é especialmente importante para conseguir tratar o outro com a dignidade que lhe é devida. O ‘outro’ de quem falo pode ser o advogado ou as partes civis, mas refiro-me em particular ao arguido num processo crime, o jogador mais frá-

gil deste jogo. Esta empatia, este ‘pôr-se no lugar do outro’, a capacidade de conseguir compreender o que o outro sente, é uma capacidade puramente humana, a máquina não a tem.

Este meu último comentário merece, porém, algumas considerações. É que cada vez começam a aparecer mais estudos a falar de uma suposta ‘empatia das máquinas’ e da criação de máquinas empáticas. Também já se fala de uma suposta ‘intuição artificial’, isto é, a intuição da IA, e também da criatividade da IA e da sua capacidade para ‘pensar fora da caixa’. No actual estado da técnica, ainda não chegámos ao ponto de a IA igualar os humanos nestas características, aparentemente exclusivamente humanas. Isto porque a IA que nós temos actualmente é a chamada IA estreita ou *narrow AI*. Ou seja, uma forma de IA que consegue desempenhar algumas funções, mas não se assemelha em nada àquilo que nós, humanos, fazemos. A ‘IA forte’ (*strong AI*), em contrapartida, será capaz de desempenhar todas as tarefas ao alcance dos humanos e tão bem quanto nós.

Quando me refiro à IA forte eu costumava usar a expressão: ‘se e quando’ aparecer a IA forte. Porém, nos últimos tempos já só digo ‘quando’. Não digo ‘se’ porque acredito que isso irá acontecer algum dia, ou seja, a IA forte será uma realidade, eventualmente já não no meu tempo de vida, mas quem sabe no dos meus filhos ou dos meus netos, ou apenas daqui a muitas gerações, mas... acontecerá. A *strong AI* poderá mudar completamente o jogo e muitos destes pontos a favor dos humanos poderão deixar de o ser. A *strong AI* será, supostamente (neste momento todas as nossas convicções não passam de meras conjecturas) aquela que conseguirá imitar-nos em todas as nossas potencialidades, ou seja, vai conseguir imitar a nossa lógica, o nosso raciocínio e todas as características humanas. Seguindo esta linha de raciocínio, irá também copiar aquelas características que até ao momento pensávamos ser puramente humanas, como a intuição, a empatia e a criatividade.

A questão das máquinas empáticas merece mais algumas considerações.

Começo por referir o óbvio: sendo a empatia exclusivamente humana, também é certo que nem todos os humanos são empáticos, logo, nem todos os juízes são empáticos. Do lado das máquinas, por outro lado, já começa a haver alguns avanços neste sentido. Não se trata ainda de empatia, pelo menos não da forma como conhecemos a empatia humana, mas poderá ser a génese de uma forma muito básica de empatia da máquina.

Deixem-me explicar o que quero dizer usando como exemplo a Alexa, a nossa já conhecida assistente pessoal digital.

Aqueles de vós que têm uma Alexa provavelmente já descobriram que tem um *feature* que se chama ‘modo sussurro’. Quando activado, cada vez que eu comunico com a Alexa em voz baixa a máquina responde de volta no mesmo tom. Não se trata, obviamente, de empatia. O que sucede é que a Alexa está a cumprir uma ordem para a qual foi programada, tal como quando eu carrego no *enter* do meu computador o cursor se move. É uma ordem que eu dou à máquina e o ‘modo sussurro’ não passa disso. Contudo, poderá chegar o dia – eventualmente com a *strong IA*, ou até ainda antes – em que nos depararemos com máquinas que são capazes de se ajustar ao nosso comportamento sem ser preciso dar a ordem. Na verdade, já dispomos de tecnologia que em boa parte o permite: temos sistemas de IA que, com base nos nossos traços faciais ou outros dados biométricos (tom de voz, por exemplo) consegue detectar as nossas emoções. Se combinarmos os diferentes métodos de análise biométrica, teremos uma IA que consegue saber o meu estado de alma. O próximo passo é fazer com que a que IA automaticamente consiga adequar o seu comportamento a esse estado de alma, o que também já se consegue fazer.

Isto não é empatia. A empatia é muito mais do que isto, bem sei, mas este poderá ser o início das ‘máquinas empáticas’ de que tanto se tem falado.

Um outro leque de argumentos que têm sido invocados para negar a substituição do juiz humano pelo ‘juiz IA’ prende-se com as especificidades do mundo jurídico. Esta argumentação prende-se com uma actividade jurídica essencial: a interpretação das normas e dos princípios jurídicos. Se eu fornecer a um sistema de IA um manual de medicina, de engenharia ou física quântica e o sistema de IA contar com um modelo de processamento de linguagem e conhecer a língua, a IA consegue perceber perfeitamente o que ali está e tomar decisões com base nesse conhecimento. Em contrapartida, se eu fornecer à IA uma lei, já duvido muito que isso aconteça, dado que a linguagem jurídica está, pela própria natureza das coisas, repleta de segundos sentidos, de apreciações éticas (e até morais) e de visões do mundo.

Digo muitas vezes aos meus alunos nas nossas aulas que para interpretar uma norma não basta saber ler e conhecer a língua na qual está escrita, porque se assim fosse não estariam a perder os melhores anos da sua vida fechados numa sala de aulas. Interpretar uma norma é muito mais do que isso, implica toda uma contextualização de valores, implica uma bagagem cultural, e não creio que neste momento tal esteja ao alcance de um sistema de IA, por mais desenvolvido que seja. A isto acresce que o elemento literal/textual não é o único, nem sequer o mais importante, na interpretação da norma jurídica. Admito até que a IA consiga fazer uma interpretação his-

tórica, basta que tenha os dados para tal. Porém, como fazer interpretações teleológicas ou sistemáticas? Se é certo que a IA não consegue neste momento fazer isto, então, ao interpretar uma lei vai estar perdida numa Torre de Babel, porque se vai ‘perder’ na linguagem especificamente jurídica.

Por exemplo, pensem nos conceitos indeterminados. O Código Civil está recheado de invocações à boa-fé e aos bons costumes. Como é que na prática o suposto ‘juiz IA’ conseguirá aplicar estes conceitos, quando é sabido que muitas vezes nem mesmo nós, os humanos, sabemos ao certo como o fazer? Tomemos como exemplo aquela que é para mim umas das normas mais difíceis de interpretar do ordenamento jurídico: a homicídio privilegiado. O que é ‘emoção violenta, compaixão, desespero ou motivo de relevante valor social ou moral’, tal como referido no artigo 133.º do Código Penal? Tenho o maior respeito pelos juízes que têm que interpretar esta norma e aplicá-la aos casos práticos da vida e penso para comigo como é que uma máquina alguma vez o conseguirá fazer. Não estou a dizer que é impossível, mas é sem dúvida um desafio.

Estas considerações conduzem-me a uma pergunta crucial: será a interpretação jurídica uma tarefa especificamente humana? Não sei qual é a resposta. Se daqui por uns anos falarmos de novo talvez já vos consiga dizer mais alguma coisa este respeito.

Um outro argumento de peso, frequentemente invocado para justificar o motivo pelo qual a IA nunca vai substituir um juiz humano, é a ausência de explicabilidade, o qual se prende com o famoso problema da caixa negra. A explicabilidade relaciona-se com uma característica da IA que é chamada de ‘obscuridade’. Nem toda a IA é obscura, ou seja, nem toda a IA é uma caixa negra. Há sistemas de IA a que chamamos de caixa branca (*white box*) porque são sistemas face aos quais se consegue perceber perfeitamente como é que a IA chegou a uma determinada conclusão. Ou seja, o agente humano consegue reconstruir o seu ‘raciocínio’. Porém, os sistemas de AI mais produtivos e eficientes que temos hoje em dia no mercado são uma autêntica caixa negra. Face a estes, conseguimos conhecer o *input*, ou seja, os dados que damos à IA (note-se que só conhecemos os dados se houver transparência, a qual pode não existir por questões relacionadas com a protecção da propriedade intelectual, segredos de negócio, vantagem competitiva), conhecemos o *output* (que é a decisão final), mas não fazemos a mínima ideia do que é que se passa entre ambos pólos. O desconhecimento quanto ao modo como a *black box* funciona afecta a todos: o cidadão comum, os que usam a IA no exercício da sua profissão e mesmo os criam os sistemas de IA. Como referido, nem toda a IA é uma caixa negra, mas a mais sofisticada IA,

aquela que verdadeiramente resolve os problemas interessantes e aquela que seria mais susceptível de substituir o juiz é, efectivamente, uma caixa negra.

A ausência de explicabilidade revela-se um problema em variados domínios, não é só no direito. Dito isto, no direito temos um desafio adicional, que não existe em muitos dos outros domínios: a imperiosa necessidade de fundamentação das decisões judiciais, tal como imposto por lei e pela Constituição. A explicação do motivo pelo qual o juiz decidiu de uma determinada forma, explicação esta que deve ser transparente e clara para quem lê a decisão, é absolutamente crucial para as partes ficarem convencidos da bondade da mesma e, por conseguinte, para garantir a paz jurídica e a tranquilidade social. É também fundamental para aqueles que não concordarem com a decisão, para que consigam recorrer e fundamentar a sua discordância. Se o ‘juiz IA’ for uma caixa negra, então, não conseguiremos ter acesso a uma fundamentação porque ninguém perceberá como é que a IA chegou àquela decisão.

A obscuridade da IA parece ser um argumento de peso para a afastar da decisão judicial e, assim, impedir a existência de um ‘juiz IA’. A questão, porém, não é assim tão simples, pois se é certo que muitos sistemas de IA são uma caixa negra, também é verdade que a verdadeira caixa negra está na nossa cabeça: o cérebro humano.

Somos humanos, e por esse simples motivo muitas das decisões que tomamos, desde as coisas mais mezinhas da vida (se visto esta camisola ou aquela) até às decisões decisivas da nossa existência, não são muitas vezes tomadas mediante passos lógicos e racionais. É frequente que tomemos uma decisão apenas porque a nossa intuição/instinto (a tal característica especificamente humana de que falei há pouco) assim o diz, e só depois, *a posteriori*, é que procuramos argumentos para justificar porque é que chegámos àquela decisão. Porém, tudo isto acontece *ex post facto*. Isto não faz de nós seres irracionais, faz de nós humanos. Há certamente decisões que são completamente lógicas, mas muitas não o são e face a estas resta-nos procurar *a posteriori* uma justificação para uma decisão tomada sem percurso lógico subjacente. Porém, se assim é, então, é possível fazer o mesmo face a decisões tomadas por uma IA caixa negra: primeiro recebemos o *outcome* e só depois fornecemos uma explicação plausível para esse *outcome*.

Aqui chegados, devo confessar que não tenho propriamente uma conclusão para as minhas considerações porque este é um trabalho em desenvolvimento. Porém, aquilo que a minha intuição me diz – mera intuição, a tal intuição tão própria dos humanos, dado que ainda não tenho factos credí-

veis para apoiar o que digo – é que a IA não vai substituir os humanos e, em particular, não vai substituir os juízes. É verdade que muito se fala das várias profissões que irão desaparecer por força da IA e certamente que assim será (desde a revolução industrial que desapareceram várias) porque há tarefas que não são dignas dos humanos. Nós, humanos, fomos feitos para coisas melhores do que meras tarefas repetitivas e, por conseguinte, se a IA conseguir desempenhar algumas tarefas que são demasiado mezinhas e insignificantes para toda a maravilhosa singularidade do ser humano, assim seja! Se a IA tiver que ler milhões e milhões de páginas de um processo para tentar detectar padrões e discrepâncias em vez de ser um humano, e o conseguir fazer com menos lapsos do que nós, perfeito, que o faça. Entretanto, vão surgir novas actividades, mais interessantes e mais desafiantes, essas sim, dignas das maravilhosas potencialidades do ser humano.

O que eu espero é que a IA nos possa complementar e se torne um suplemento precioso para as nossas funções, inclusivamente para nos tornarmos melhores humanos. Não humanos ao serviço das máquinas, mas humanos com mais capacidades e capazes de fazer coisas ainda melhores.

Claro que posso estar enganada. Não sou o Nostradamus e nem sequer ele adivinhou tudo. O futuro pode vir a desenrolar-se de forma totalmente contrária às minhas previsões, que talvez não passem de meras aspirações de alguém profundamente apaixonado pela ciência, pela tecnologia e pelo futuro. Admito igualmente esta hipótese.

Em última instância, tudo depende de nós e do quadro tecnológico, jurídico e ético que venhamos a construir para a IA. Depende de nós saber se vamos viver no pior dos tempos, na época da escuridão e da descrença, onde a humanidade se encontre ao serviço da máquina; ou se, ao invés, vamos estar no melhor dos tempos, nos tempos da esperança, da luz e da sabedoria, em que a máquina estará ao serviço da humanidade para nos tornarmos melhores humanos. •

Muito obrigada pela vossa atenção.





BREVES REFLEXÕES SOBRE ÉTICA E A SUA APLICAÇÃO À INTELIGÊNCIA ARTIFICIAL



Nuno Lima da Luz¹
Advogado Associado Sênior, Cuatrecasas

Clique para assistir.



RESUMO: O documento apresenta uma reflexão sobre os princípios éticos e a regulação da inteligência artificial (IA), abordando os desafios e as oportunidades que esta tecnologia traz para a sociedade e para o direito. O autor explora as teorias de filósofos jurídicos como Rawls e Dworkin, e aplica os seus conceitos de justiça, equidade e integridade à construção e uso de modelos de IA. O autor também discute o estado atual da regulação da IA em diferentes jurisdições, destacando as diferentes abordagens e soluções adotadas. Por fim, o autor propõe algumas soluções possíveis para enfrentar os desafios éticos e legais da IA, como a criação de padrões e diretrizes internacionais, a colaboração entre as partes interessadas, a promoção da educação e consciencialização sobre a IA, a implementação de sistemas de auditoria e certificação e a atualização constante da regulação. O autor defende que a regulação da IA é uma questão de equilíbrio entre a promoção da inovação e a garantia da proteção dos direitos, liberdades e garantias dos utilizadores da tecnologia.

PALAVRAS-CHAVE: ética, regulação, inteligência artificial, justiça, equidade, integridade, inovação, direitos.

1 | Introdução

Como juristas, somos especialmente conscientes da necessidade de equilibrar a inovação com a proteção dos direitos e garantias fundamentais dos cidadãos. Ora, a Inteligência Artificial já demonstrou ter o potencial de transformar a nossa sociedade de maneiras revolucionárias, trazendo benefícios significativos em termos de eficiência, produtividade e qualidade de vida.

¹ Advogado Sênior na área de Tecnologia e Digital Business da Cuatrecasas, Gonçalves Pereira & Associados, Sociedade de Advogados, SP, RL

No entanto, juntamente com as oportunidades, surgem também desafios significativos relacionados à ética, à justiça e à proteção dos direitos e garantias dos cidadãos.

Ao longo deste breve artigo, vamos explorar os princípios da ética da IA, centrando-nos na justiça, responsabilidade, transparência e segurança, e discutir o estado atual da legislação relacionada com a IA em diferentes jurisdições. O objetivo é deixar algumas notas que nos permitam pensar, refletir e idealizar algumas estratégias para encontrar o equilíbrio certo entre a promoção da inovação em IA e a garantia da proteção dos direitos, liberdades e garantias dos utilizadores desta tecnologia.

2 | Princípios éticos em Direito e a sua aplicação à tecnologia

A interseção entre ética, Direito e tecnologia é um campo de exploração rica, especialmente à luz das teorias de filósofos jurídicos conceituados como Rawls ou Dworkin. John Rawls, através de sua teoria da “justiça como equidade”, propôs que um sistema justo é aquele que beneficia os menos favorecidos na sociedade². Ele argumentou que, numa “posição original” hipotética, onde os indivíduos não têm conhecimento de sua posição social ou identidade pessoal, eles escolheriam princípios de justiça que beneficiariam todos, incluindo os mais desfavorecidos. Esta abordagem pode ser aplicada ao mundo da tecnologia, onde questões sobre o acesso equitativo a recursos tecnológicos e proteção de dados estão em destaque.

A discussão ética no direito também é enriquecida pela contribuição de outros filósofos como Ronald Dworkin, que enfatiza a importância da integridade no Direito. Dworkin abordou a importância da integridade no Direito na sua obra “Law’s Empire” (O Império do Direito), apresentando-nos a sua teoria do Direito como integridade, onde argumenta que os juízes devem interpretar a lei em termos de princípios morais consistentes, especialmente justiça e equidade. A integridade, segundo Dworkin, é manifestada quando o estado exerce seu poder coercitivo de acordo com um conjunto coerente de princípios de justiça, justiça e devido processo legal.

Em “Law’s Empire”, Dworkin avança com a ideia de que, sob a perspectiva da integridade, os juízes são obrigados a decidir casos de uma maneira que esteja em coerência com os princípios morais fundamentais, promovendo assim a justiça, a equidade e o devido processo legal, tendo esta obra vindo a ser bastante influente na moldagem de teorias contemporâneas sobre a natureza do direito.

² Rawls, J. (1971). “A Theory of Justice”. Harvard University Press.

Na citada obra, propõe Dworkin que: “A integridade é a chave para entender a prática jurídica. O império do Direito é definido pela atitude, não pelo território ou poder ou processo.”³ Esta abordagem, ou proposta, tem vindo a influenciar até aos dias de hoje as diversas discussões éticas e legais no contexto da tecnologia, onde a integridade na interpretação e aplicação do direito é crucial para abordar desafios emergentes como privacidade, discriminação algorítmica e governança de Inteligência Artificial (IA).

No âmbito da tecnologia, as reflexões éticas e legais são cruciais para navegar em questões emergentes como a IA e a autonomia algorítmica, onde a aplicação de princípios éticos fundamentais no desenvolvimento e implementação de IA pode ajudar a construir sistemas que respeitam os direitos humanos, a privacidade e a justiça social. Por exemplo, a regulação ética da IA pode garantir que os sistemas algorítmicos sejam transparentes, justos e não discriminatórios, alinhando-se com os ideais de justiça e equidade propostos por filósofos como Rawls e Dworkin.

Além disso, a filosofia do Direito oferece-nos as ferramentas necessárias para avaliar como os quadros legais existentes, e os vindouros, podem moldar o desenvolvimento tecnológico de maneira ética e justa. A análise crítica e contínua das implicações éticas e legais da tecnologia é essencial para garantir que as inovações tecnológicas sejam conduzidas de maneira responsável e alinhadas com os valores sociais fundamentais. A interação entre ética, Direito e tecnologia continuará a ser uma área chave de investigação e debate, à medida que a sociedade procura harmonizar os avanços tecnológicos com os princípios éticos e legais fundamentais.

A ascensão vertiginosa da inteligência artificial nos últimos anos tem sido um testemunho da capacidade humana de inovar e transformar o tecido da nossa sociedade. Com aplicações que vão desde o diagnóstico médico até a tomada de decisões financeiras, a IA tem vindo a redefinir os contornos da nossa interação com a tecnologia e até entre humanos. No entanto, à medida que avançamos nessa jornada, enfrentamos desafios éticos e regulatórios que exigem uma análise cuidadosa.

Sendo a IA uma das áreas de maior desenvolvimento e inovação na atualidade, esta tem vindo a oferecer soluções para uma ampla variedade de desafios em muitos setores, como saúde, finanças, transporte, comunicação, entre outros. No entanto, essa tecnologia também traz consigo desafios significativos, exigindo que todos os intervenientes no seu desenvolvimento e colocação em mercado, desde os programadores até ao legislador que enquadra e regula este tipo de atividades, estejam cientes desses desafios e

³ Dworkin, R. (1986). “Law’s Empire”. Harvard University Press.

tomem medidas para lidar com eles de forma adequada.

A IA, com sua capacidade de tomar decisões que anteriormente eram do domínio exclusivo dos seres humanos, traz à tona a questão da transparência. Em muitas jurisdições, a transparência emergiu como um pilar ético na regulação da IA. Se um algoritmo de IA é usado, por exemplo, para determinar a elegibilidade de um paciente para um tratamento médico específico, é imperativo que o processo de tomada de decisão seja transparente. Os pacientes têm o direito de entender como as decisões que afetam diretamente suas vidas são tomadas⁴.

Ora, a ética, na sua essência, é um ramo da filosofia que estuda os princípios e valores morais que orientam o comportamento humano na sociedade, bem como os princípios que governam o comportamento humano, servindo como uma bússola moral para as nossas ações. No domínio jurídico, a ética desempenha um papel fundamental na condução de profissionais e na construção de leis e regulamentos que promovam a justiça, a igualdade e a dignidade humana. Os profissionais do Direito, ao interpretarem e aplicarem leis, devem ser guiados por princípios éticos que garantam justiça, igualdade e dignidade humana, e também, esses mesmos princípios éticos, devem estar presentes em outras dimensões da vida em sociedade, nomeadamente no campo da tecnologia.

Embora a regulação da IA seja um tópico em evolução, um dos principais princípios éticos que emergiram em muitas jurisdições é o princípio da transparência. Isso significa que as empresas que desenvolvem sistemas de IA devem ser transparentes sobre como esses sistemas funcionam e como eles tomam decisões, sendo essa transparência particularmente importante em setores que afetam a vida e os direitos dos indivíduos, como saúde, finanças e justiça.

Da mesma forma, a responsabilidade é outro princípio ético fundamental no desenvolvimento da IA, levantando-se algumas questões pertinentes neste domínio: Se um sistema de IA comete um erro, quem é responsável? Os programadores que desenvolveram o algoritmo? A empresa que implementou a solução? Ou o próprio algoritmo? Estas questões merecem a nossa atenção por forma a garantir que a tecnologia serve a humanidade de maneira justa e ética, e não sermos subjugados à mesma por força da sua utilização⁵. De um modo geral, significa que quem desenvolve algoritmos de IA deva passar num crivo de responsabilidade pelas decisões que os sistemas

⁴ Bostrom, N., & Yudkowsky, E. (2014). "The Ethics of Artificial Intelligence". Cambridge Handbook of Artificial Intelligence.

⁵ Vincent, J. (2017). "Who's to blame when a machine botches your surgery?". The Verge.

por si desenvolvidos tomam e pelos impactos que essas decisões algorítmicas possam ter na vida das pessoas. Esse princípio de responsabilidade também implica que quem desenvolve estes algoritmos deva fornecer soluções para possíveis erros ou danos causados pelos seus sistemas.

Outro princípio ético essencial é o princípio da justiça, que implica, no nosso caso de estudo, que os sistemas de IA devam ser justos e imparciais em relação a todas as pessoas, independentemente de sua raça, género, religião, orientação sexual ou outras características intrinsecamente humanas. Este princípio será particularmente importante no sector judiciário, onde os sistemas de IA podem ser usados para tomar decisões que afetam a liberdade e os direitos dos cidadãos e que, não sendo utilizados devidamente, poderão ter consequências altamente gravosas.

Mas, além dos princípios éticos fundamentais acima referidos, existem outros que também são da maior importância e relevância quando discutimos o tópico da regulação da IA, como sejam a privacidade e segurança.

A questão da privacidade é particularmente importante em setores que lidam com dados sensíveis, como saúde e finanças, e deve ser protegida através de políticas e tecnologias adequadas. Já a segurança é fundamental para garantir que os sistemas de IA não sejam comprometidos ou usados para fins maliciosos. Estes dois princípios, aos dias de hoje e com o advento da globalização tecnológica, adquirem toda uma nova dinâmica e acrescidas preocupações.

Assim, aplicando os referidos princípios éticos à tecnologia da IA, é importante que o legislador coopere muito de perto com quem desenvolve e coloca em mercado estes algoritmos, e outros profissionais do setor, para garantir que a regulação da IA seja eficaz e equilibrada, através da criação de comités especializados para fornecer orientação sobre questões éticas na construção dos modelos algorítmicos, bem como a criação de fóruns para discussão e colaboração entre as diferentes partes interessadas.

Acresce também ao exposto acima a importante necessidade de que a legislação aplicável seja flexível e adaptável às mudanças tecnológicas em curso, idealmente tecnologicamente agnóstica, de modo a garantir que se possam manter atualizadas à medida que novas questões éticas e desafios legais surgem. A regulação também deve ser equilibrada, para garantir que não se impede ou limita o desenvolvimento e a inovação na área, mas, ao mesmo tempo, que se consiga proteger os direitos, liberdades e garantias dos utilizadores da tecnologia.

Por fim, é importante destacar que a aplicação dos princípios éticos na regulação da IA é crucial para garantir que essa tecnologia seja usada de maneira responsável e justa. Com a rápida evolução da IA e seu impacto cada vez maior na vida dos indivíduos e da sociedade como um todo, a regulação adequada e equilibrada da IA é mais importante do que nunca. Podemos assim garantir que a IA é usada para o bem-estar humano e para o progresso social e económico, evitando que a tecnologia seja usada de maneira prejudicial ou injusta.

Consideradas as ideias até aqui expostas, torna-se por demais evidente que, a aplicação de princípios éticos na regulação da IA é uma questão crucial na atualidade. É essencial que os participantes no desenvolvimento desta tecnologia, legislador e outros profissionais do setor trabalhem juntos para garantir que a regulação seja eficaz, flexível e equilibrada, para que possamos garantir que a tecnologia da IA seja usada de maneira responsável e justa, beneficiando a sociedade como um todo.

3 | Considerações práticas sobre princípios éticos e a sua aplicação em modelos de IA

À medida que a inteligência artificial se torna mais sofisticada e integrada aos sistemas que moldam a nossa sociedade, a sua capacidade de tomar decisões complexas aumenta exponencialmente. Estas decisões não são meras escolhas binárias, mas avaliações ponderadas baseadas em vastos conjuntos de dados que vão além da capacidade humana de processar.

Esta automação decisória da IA pode ser vista em várias áreas, desde sistemas de recomendação que sugerem o próximo filme a assistir, até algoritmos que decidem quem deve receber crédito ou como os recursos médicos devem ser alocados em situações de emergência. Embora essas decisões automatizadas possam trazer eficiência e objetividade, elas também levantam preocupações significativas sobre justiça, equidade e direitos humanos⁶.

Um dos principais desafios é o risco de viés nos sistemas de IA. Os algoritmos são tão bons quanto os dados em que são treinados. Se esses dados refletirem preconceitos históricos ou atuais, a IA pode perpetuar ou até exacerbar essas injustiças⁷. Por exemplo, tem havido casos documentados onde sistemas de reconhecimento facial exibem taxas de erro mais altas para grupos étnicos específicos, levantando sérias questões sobre discriminação

⁶ O'Neil, C. (2016). "Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy". Crown.

⁷ Buolamwini, J., & Gebru, T. (2018). "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification". Proceedings of Machine Learning Research.

e imparcialidade⁸.

Além disso, a tomada de decisão automatizada carece de alguma empatia que as máquinas ou algoritmos ainda não conseguem emular de forma conveniente ou satisfatória. Enquanto os humanos podem considerar circunstâncias atenuantes ou nuances emocionais ao tomar decisões, baseadas no reconhecimento empírico de traços fisiológicos apurados ao longo de milhares de anos, os algoritmos fazem julgamentos baseados estritamente em dados e padrões previamente reconhecidos. Isso pode resultar em decisões que, embora lógicas do ponto de vista do algoritmo, pareçam frias ou insensíveis aos seres humanos afetados pelas mesmas.

Ora, dada a crescente dependência da IA na tomada de decisões que afetam a vida dos cidadãos, é imperativo que haja um escrutínio rigoroso e uma regulação eficaz para garantir que esses sistemas operem de maneira justa e ética pois, como já vimos anteriormente, a inteligência artificial tem sido amplamente adotada em diversas áreas e setores críticos à nossa vida em sociedade, como saúde, finanças, transportes e educação, para automatizar processos e melhorar a tomada de decisões.

No entanto, o crescente uso de IA levanta questões que importa abordar, relativamente ao impacto significativo na vida dos cidadãos. Neste contexto, é crucial entender os impactos da IA na vida dos cidadãos e abordar os desafios éticos e legais associados ao seu uso, pelo que referimos, a título de exemplo, alguns dos principais impactos e desafios que a utilização desta tecnologia acarreta:

1. **Discriminação e viés algorítmico:** Os algoritmos de IA são treinados com base em dados históricos e podem, portanto, perpetuar e agravar os vieses e desigualdades existentes na sociedade, ainda que seja puramente factual ou estatístico. Isso pode levar a decisões automatizadas que discriminam injustamente certos grupos ou indivíduos com base em atributos protegidos, como raça, género ou idade. Para abordar esse desafio, é fundamental garantir que os algoritmos de IA sejam projetados e treinados de maneira justa e imparcial, bem como implementar mecanismos de monitorização e auditoria para identificar e corrigir esses vieses.

⁸ Raji, I. D., & Buolamwini, J. (2019). "Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products". AAAI/ACM Conference on AI, Ethics, and Society.

- 2. Transparência e explicabilidade:** A tomada de decisão baseada em IA pode ser opaca e difícil de entender para os cidadãos afetados, levantando preocupações sobre a responsabilidade e a justiça, no caso de se verificar algum dano efetivo baseado num mau julgamento ou aplicação prática da tecnologia. É essencial garantir que os processos e algoritmos de tomada de decisão sejam transparentes e explicáveis, permitindo que os cidadãos compreendam como e por que as decisões são tomadas e possam contestá-las, se necessário. Isso pode envolver o desenvolvimento de técnicas de IA explicável e a adoção de padrões de transparência e divulgação de informações.
- 3. Privacidade e proteção de dados:** A IA geralmente depende do processamento de grandes volumes de dados pessoais e sensíveis, levantando preocupações sobre a privacidade e a segurança dos dados. Torna-se, pois, crucial garantir que os dados sejam recolhidos, armazenados e tratados de acordo com as leis e regulamentos de proteção de dados, como o Regulamento Geral sobre a Proteção de Dados, e implementar medidas para garantir a anonimização e a minimização dos dados sempre que possível. Além disso, é importante informar os cidadãos sobre como os seus dados são usados e garantir que eles controlem a recolha e uso dos seus dados pessoais.
- 4. Responsabilidade:** A tomada de decisões baseadas em IA pode levantar questões sobre quem terá a efetiva responsabilidade civil por decisões que resultam em consequências negativas para os cidadãos. É essencial garantir clareza e previsibilidade sobre quem é responsável por essas decisões automatizadas e que existam mecanismos claros para remediar danos causados por decisões incorretas ou prejudiciais. A solução tem passado por incluir o desenvolvimento de regimes de responsabilidade civil ou a criação de órgãos reguladores para supervisionar e fiscalizar o uso da IA, que abordaremos mais à frente.
- 5. Justiça e equidade:** A tomada de decisão automatizada pode afetar desproporcionalmente certos grupos ou indivíduos. É, assim, crucial garantir que as decisões automatizadas sejam projetadas e implementadas de maneira a promover a justiça e a igualdade, e não para

reforçar desigualdades ou preconceitos existentes. Algumas medidas mitigatórias poderão envolver a adoção de mecanismos que garantam uma efetiva diversidade e inclusão nos dados e algoritmos utilizados, bem como a implementação de mecanismos de correção para identificar e mitigar possíveis impactos negativos.

A construção de modelos de inteligência artificial não é, pois, apenas um exercício técnico; é também uma atividade profundamente ética. Cada decisão tomada durante o desenho, treino e implementação de um modelo algorítmico pode ter implicações éticas, com um impacto real na vida dos indivíduos e da sociedade em geral.

A construção de modelos de inteligência artificial envolve uma série de decisões complexas, desde a escolha dos algoritmos e dados de treino até a implementação e monitorização do modelo aplicado. Em cada uma dessas etapas, é possível aplicar os princípios éticos em direito à tecnologia para garantir que a IA é desenvolvida e usada de maneira responsável e justa. Os desafios anteriormente referidos podem ser mitigados através da aplicação de alguns requisitos ou princípios, conforme exemplificamos abaixo:

1. **Escolha de algoritmos:** Os algoritmos de IA variam na sua capacidade de tomar decisões imparciais e justas. Alguns algoritmos podem perpetuar preconceitos e discriminação, enquanto outros são projetados especificamente para evitar vieses. Ao escolher os algoritmos para uso em modelos de IA, é crucial considerar não apenas a precisão e eficiência do algoritmo, mas também a sua capacidade de tomar decisões justas e imparciais.
2. **Seleção de dados de treino:** Os dados usados utilizados para treinar modelos de IA podem conter vieses e desigualdades, refletindo as desigualdades e preconceitos existentes na sociedade. É importante garantir que esses dados sejam garantidamente representativos e imparciais, no que toca à representação da realidade. Podemos mitigar o problema através da implementação de mecanismos de correção para identificar e mitigar vieses nos dados de treino.
3. **Avaliação de desempenho:** A avaliação do desempenho de modelos de IA pode ser complexa, envolvendo a análise de várias mé-

tricas, como precisão, sensibilidade e especificidade. No entanto, é importante não nos focarmos apenas em métricas de desempenho, mas também considerar questões éticas, como a justiça e equidade das decisões tomadas pelo modelo. Uma possível solução seria através da adoção de métricas de desempenho ético, como a diversidade e representatividade dos resultados do modelo.

4. **Implementação e monitorização do modelo:** Uma vez implementado o modelo de IA, é importante fazer uma monitorização contínua para garantir que está a funcionar conforme o esperado e que é utilizado de maneira responsável. Será, pois, relevante a implementação de mecanismos de monitorização e auditoria para detetar vieses e desigualdades, bem como a criação de sistemas de responsabilização para remediar danos causados pelo modelo.
5. **Acesso e transparência:** A transparência e o acesso aos modelos de IA podem ajudar a promover a responsabilidade e a justiça na tomada de decisões. É importante garantir que os modelos de IA sejam transparentes e explicáveis, permitindo que os cidadãos compreendam como e por que as decisões são tomadas e possam contestá-las, se necessário. Essa transparência pode ser verificada através da implementação de mecanismos de acesso e divulgação de informações sobre os modelos de IA.

Em suma, e como temos vindo a aflorar até aqui, com a aplicação de princípios éticos na construção de modelos de IA, é possível garantir que a IA seja usada de maneira responsável e justa, evitando discriminação e desigualdades injustas na tomada de decisões automatizadas. Ainda assim é importante destacar que a aplicação desses mesmos princípios não é uma solução única e perfeita para os desafios éticos e legais associados à IA, pois isso exige um equilíbrio delicado entre a promoção da inovação e a garantia da proteção dos direitos, liberdades e garantias dos utilizadores da tecnologia.

4 | Enquadramento legal relativamente à Inteligência Artificial

A rápida evolução da IA trouxe desafios significativos para os legisladores e reguladores em todo o mundo. Embora a tecnologia prometa benefícios substanciais em termos de eficiência e inovação, também levanta questões

sobre privacidade, segurança, responsabilidade e equidade. Como resultado, diferentes países adotaram abordagens variadas para enquadrar a IA legalmente, refletindo suas próprias prioridades, valores e contextos culturais.

1. **União Europeia (UE):** A UE tem sido pioneira na regulação da tecnologia digital, e a IA não é exceção. A Comissão Europeia propôs regulamentos que enfatizam a transparência, a responsabilidade e os direitos fundamentais. O Regulamento Geral de Proteção de Dados (GDPR) também tem implicações para a IA, especialmente no que diz respeito ao processamento de dados pessoais e ao direito de explicação⁹. A União Europeia (UE), por exemplo, introduziu o Regulamento de Inteligência Artificial em abril de 2021. A regulação propõe uma abordagem baseada em riscos para orientar o uso da IA tanto no setor público quanto privado. O regulamento define três categorias de risco, incluindo aplicativos de risco inaceitável, aplicativos de alto risco e aplicativos não explicitamente proibidos. Embora a legislação ainda esteja em análise no Parlamento Europeu, este Regulamento proíbe o uso da IA em serviços críticos que possam ameaçar os meios de subsistência ou incentivar comportamentos destrutivos, enquanto permite o uso da tecnologia em outros setores sensíveis, como a saúde, com verificação máxima de segurança e eficácia pelos reguladores. O Regulamento foi recentemente finalizado pelo Parlamento Europeu e pelo Conselho Europeu, terá certamente o potencial de se tornar a norma global de facto para a Inteligência Artificial, também volta a enfatizar a proteção dos direitos humanos. Para além de proibir algoritmos que tenham o potencial de violar estes direitos, também obriga a uma nova avaliação de impacto sobre os direitos humanos fundamentais, bem como a um reforço das avaliações de impacto sobre a proteção de dados. Este Regulamento de Inteligência Artificial já está em vigor e torna-se plenamente aplicável em meados de 2026.

2. **Estados Unidos:** Nos EUA, a abordagem tem sido mais fragmentada, com diferentes estados propondo ou implementando a sua própria legislação de IA. No entanto, em nível federal, tem havido discussões sobre a criação de uma estrutura regulatória mais coesa, focada na

⁹ European Commission. (2021). “Proposal for a Regulation on a European approach for Artificial Intelligence”. Official Journal of the European Union.

promoção da inovação enquanto se protegem os direitos dos cidadãos¹⁰. No Estados Unidos, em 4 de outubro de 2022, o Gabinete de Política Científica e Tecnológica da Casa Branca publicou um projeto para o desenvolvimento, utilização e implantação de sistemas automatizados (projeto para uma declaração de direitos da IA). O projeto é – ao contrário do projeto de lei da UE sobre a IA – não vinculativo e enumera cinco princípios que se destinam a minimizar os danos potenciais dos sistemas de IA. Em 18 de Agosto de 2022, o Instituto Nacional de Normas e Tecnologia (NIST) publicou o segundo projeto do seu Quadro de Gestão de Riscos da IA para comentários. A versão original data de março de 2022 e baseia-se num documento conceptual de dezembro de 2021. O Quadro de Gestão de Riscos da IA destina-se a ajudar as empresas que desenvolvem ou implementam sistemas de IA a avaliar e gerir os riscos associados a estas tecnologias. Consiste em orientações e recomendações voluntárias, pelo que também não é vinculativo e não deve ser explicitamente entendido como um regulamento.

3. **China:** A China adotou uma abordagem pragmática, equilibrando o desejo de ser líder mundial em tecnologia de IA com a necessidade de abordar questões de segurança e ética. A China introduziu diretrizes que enfatizam a pesquisa responsável e o desenvolvimento de IA, bem como a proteção da privacidade dos dados¹¹. Na China, o governo tem vindo a regular ativamente a IA desde 2017, com a criação do “Plano de Desenvolvimento da Inteligência Artificial da Próxima Geração”. Em 2021, foram publicadas orientações éticas, seguidas da publicação de duas leis em 2022 relacionadas com aplicações específicas de IA. As “Disposições sobre Algoritmos”, que abordam o abuso de sistemas de recomendação algorítmica, estão em vigor desde março de 2023, enquanto o “Projeto de Disposições sobre Síntese Profunda” (*deep synthesis ou deep fakes*), destinado a regular as tecnologias de síntese profunda para combater as falsificações profundas (ou *deep fakes*), ainda se encontra em fase de projeto. De referir que a Administração do Ciberespaço da China realizou recentemente uma consulta sobre o projeto de Medidas Administrativas para Serviços de Inteligência Artificial Generativa. Foi também publicada legislação regional, com Shenzhen e Xangai a publicarem regula-

¹⁰ U.S. Congress. (2019). “Artificial Intelligence Initiative Act”. 116th Congress.

¹¹ China’s State Council. (2017). “New Generation Artificial Intelligence Development Plan”.

mentos sobre a promoção do desenvolvimento da indústria da IA. Estes regulamentos abrangem a gestão de conteúdos, a marcação ou rotulagem, a transparência, a proteção de dados, as práticas leais e as sanções por incumprimento. Os prestadores de serviços de IA devem tomar medidas para impedir a produção de informações falsas e evitar conteúdos nocivos.

4. **Outros Países:** Muitos outros países estão em diferentes estágios de desenvolvimento de suas próprias estruturas regulatórias para IA. Por exemplo, a Índia tem focado na ética e na equidade, enquanto o Japão tem priorizado a promoção da inovação e a cooperação internacional¹².

Resumindo, a regulação da IA é uma preocupação global crescente e os governos procuram maneiras de equilibrar os benefícios da tecnologia com os riscos e desafios associados. A abordagem de regulação varia de país para país, com alguns países que proíbem *tout court* o uso de IA em setores críticos, enquanto outros optam por regulamentar casos de uso específicos. A UE e o Canadá adotaram uma abordagem baseada em riscos para orientar o uso da IA, enquanto os EUA e a China ainda desenvolvem regulação a nível federal. A regulação de IA provavelmente continuará a evoluir à medida que a tecnologia se torna mais amplamente utilizada e novos desafios e riscos surgem.

5 | Quais serão os próximos passos?

O futuro da IA é repleto de oportunidades e desafios e, embora tenhamos feito progressos significativos na última década, ainda estamos no início de compreender o verdadeiro potencial e as implicações da IA. Para garantir que esta tecnologia beneficie a humanidade e preserve os valores éticos e morais, é crucial considerar os caminhos e soluções possíveis para os próximos anos.

Embora a regulação da IA esteja em evolução em todo o mundo, ainda há desafios significativos a serem enfrentados na construção de modelos de IA responsáveis e justos, havendo, no entanto, algumas soluções potenciais que podem ajudar a enfrentar esses desafios e promover o desenvolvimento responsável da tecnologia.

¹² Global Partnership on Artificial Intelligence (GPAD). (2020). “National AI Strategies and Policies: An Overview”.

Uma abordagem possível é a criação de padrões e diretivas internacionais para a regulação da IA e que poderia passar pela criação de um organismo internacional responsável por definir e promover a adoção de padrões éticos para a IA, bem como por estabelecer diretivas para a regulação da IA em diferentes jurisdições, como aliás já temos vindo a observar.

Além disso, a colaboração entre diferentes partes interessadas, incluindo programadores de IA, reguladores e profissionais do Direito, pode ajudar a garantir que a regulação da IA seja equilibrada e eficaz. A promoção da educação e conscientização sobre a IA é, por isso, também importante para garantir que quem desenvolve este tipo de algoritmos esteja ciente dos desafios éticos envolvidos na construção e uso desses modelos, através da criação de disciplinas e cadeiras focadas em ética, nos cursos superiores de engenharia e programação, bem como a realização de campanhas de conscientização pública sobre os impactos da IA na sociedade.

Também a implementação de sistemas de auditoria e certificação pode ajudar a garantir que os modelos de IA sejam construídos e usados de maneira responsável e justa, ou pelo menos mantidos como tal. Através da implementação de sistemas de auditoria independentes para avaliar a precisão, transparência e justiça dos modelos de IA, bem como a criação de programas de certificação para empresas que seguem as melhores práticas na construção e uso de modelos de IA, podemos atingir o objetivo pretendido.

Em jeito de sumário, a regulação da IA é um desafio complexo que exige a colaboração de diferentes partes interessadas para promover o desenvolvimento responsável da tecnologia. A criação de padrões e diretrizes internacionais, a colaboração entre diferentes partes interessadas, a promoção da educação e conscientização sobre a IA, a implementação de sistemas de auditoria e certificação e a atualização constante da regulação são algumas soluções potenciais que podem ajudar a enfrentar os desafios éticos e legais da IA. À medida que a tecnologia da IA continua a evoluir e se tornar mais integrada nas nossas vidas e atividades diárias, é crucial que as soluções para a regulação da IA sejam implementadas para garantir que a tecnologia seja usada de maneira responsável e justa, protegendo os direitos e a segurança dos utilizadores da tecnologia.

No entanto, é importante frisar que a regulação da IA não deve ser vista como uma barreira ao desenvolvimento da tecnologia, mas sim como um meio de promover a sua inovação e progresso sustentável. Uma regulação equilibrada e eficaz pode permitir um mercado justo e seguro para quem

desenvolve este tipo de soluções e ao mesmo tempo garantir e proteger os direitos e a segurança dos cidadãos. A regulação da IA será, então, sempre uma questão de equilíbrio entre a promoção da inovação e a garantia da proteção dos direitos, liberdades e garantias dos utilizadores da tecnologia.

E, nessa medida, como síntese do que já foi abordado até então, deixamos alguns tópicos que devem ser levados em linha de conta num futuro próximo:

1. **Educação e Formação:** À medida que a IA se torna mais prevalente, é vital educar tanto os profissionais que trabalham diretamente com esta tecnologia quanto o público em geral. A compreensão dos princípios básicos da IA e dos seus impactos potenciais pode ajudar a informar decisões éticas e responsáveis¹³.
2. **Padrões Éticos Universais:** A criação de padrões éticos universais para o desenvolvimento e aplicação da IA pode fornecer uma base sólida para a comunidade global. Estes padrões podem servir como um guia, garantindo que a IA seja utilizada de forma a respeitar os direitos e a dignidade do ser humano¹⁴.
3. **Transparência e Explicabilidade:** Continuar a enfatizar a importância da transparência na construção de modelos de IA é crucial. Além disso, os modelos devem ser explicáveis, permitindo que os utilizadores compreendam como e por que razão certas decisões são tomadas¹⁵.
4. **Abordagem Centrada no Humano:** Em vez de ver a IA como uma ferramenta que opera isoladamente, é crucial adotar uma abordagem centrada no ser humano. Isso significa que a IA deve ser desenvolvida e aplicada de maneira a complementar e ampliar as capacidades humanas, em vez de substituí-las¹⁶.
5. **Cooperação Internacional:** Dada a natureza global da tecnologia e dos desafios que a IA apresenta, a cooperação internacional é essencial. Países e organizações devem trabalhar juntos para compar-

¹³ Russell, S. (2019). "Human Compatible: Artificial Intelligence and the Problem of Control". Viking.

¹⁴ IEEE. (2019). "Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems".

¹⁵ Wachter, S., Mittelstadt, B., & Floridi, L. (2017). "Transparent, explainable, and accountable AI for robotics". Science Robotics.

¹⁶ Brynjolfsson, E., & McAfee, A. (2014). "The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies". W. W. Norton & Company.

tilhar conhecimentos, estabelecer normas e garantir que a IA seja usada para o benefício de todos¹⁷.

6 | Conclusão

Conforme explorado nesta brevíssima análise, a IA é uma tecnologia em rápida evolução que traz consigo desafios significativos em relação à ética e à regulação. Embora a IA ofereça muitas oportunidades para inovação e progresso em diferentes setores, é crucial que essa tecnologia seja usada de maneira responsável e justa.

Para enfrentar esses desafios, a componente regulatória tem vindo a ser desenvolvida em todo o mundo, com vários princípios éticos comuns. Esses princípios incluem transparência, responsabilidade, justiça, privacidade e segurança, e são fundamentais para garantir que a IA seja usada de maneira responsável. No entanto, a implementação desses princípios pode variar significativamente entre diferentes países e organizações, e ainda há desafios significativos a serem enfrentados na regulação da IA. Para superar esses desafios, é importante que haja colaboração entre as diferentes partes interessadas, incluindo os criadores dos algoritmos, reguladores, profissionais do Direito e legislador, para encontrar soluções eficazes e equilibradas para a regulação da IA.

Algumas soluções potenciais incluem a criação de padrões e diretivas internacionais para a regulação da IA, a colaboração entre diferentes partes interessadas, a promoção da educação e consciencialização sobre a IA, a implementação de sistemas de auditoria e certificação e a atualização constante da regulação. Em última análise, regular devidamente é uma questão de equilíbrio entre a promoção da inovação e a garantia da proteção dos direitos, liberdades e garantias. Com o trabalho conjunto e colaboração de diferentes partes interessadas, é possível encontrar o equilíbrio certo e garantir que a IA seja usada de maneira responsável e justa para o benefício de todos.

À medida que a tecnologia da IA continua a desenvolver-se e a integrar-se nas nossas vidas, é crucial que continuemos a acompanhar estas evoluções e que se coopere para um uso saudável da tecnologia. •



¹⁷ United Nations. (2020). “Roadmap for Digital Cooperation”.

TROCA DE
EXPERIÊNCIAS
COM PAÍSES
INTEGRANTES
DA CIMEIRA
JUDICIAL IBERO-
-AMERICANA

INTELIGENCIA ARTIFICIAL (I.A) Y SISTEMAS DE JUSTICIA: TRES TEMÁTICAS Y ALGUNAS APLICACIONES CONCRETAS EN EL PODER JUDICIAL DE ESPAÑA



Juan Martínez Moyaz

Magistrado y Vocal del consejo general del Poder judicial de España

Miembro de la Comisión de Coordinación

y Seguimiento de la Cumbre Judicial Iberoamericana

Clique para assistir.



RESUMEN: El impacto de la Inteligencia Artificial en el sistema de justicia ni debe sobrestimarse ni tampoco subestimarse. Plantea variados problemas: desde la necesidad de plantearse su codificación normativa, pasando por la influencia en la modernización de los sistemas judiciales y la necesidad de preservar el factor humano como última ratio. La IA aplicada a la Administración de Justicia debe ser abordada en el contexto de los estándares del Estado de Derecho lo que condicionará los contenidos y límites legales y éticos. La IA también está llamada a crear un nuevo paradigma en la vertebración técnica del Derecho. En el caso de España, el Centro de Documentación Judicial (CENDOJ), órgano técnico dependiente del Consejo General del Poder Judicial de España, ha desarrollado varias utilidades basadas en Inteligencia Artificial con la finalidad de dar respuesta a las necesidades de los jueces en el ámbito del conocimiento judicial.

ABSTRACT: The impact of Artificial Intelligence on the justice system should neither be overestimated nor underestimated. It raises various problems: from the need to consider its normative codification, through the influence on the modernization of judicial systems and the need to preserve the human factor as the last ratio. AI applied to the Administration of Justice must be addressed in the context of the standards of the Rule of Law, which will condition the legal and ethical contents and limits. AI is also called to create a new paradigm in the technical structure of Law. In the case of Spain, the Judicial Documentation Center (CENDOJ), a technical body dependent on the General Council of the Judiciary of Spain, has developed several utilities based on Artificial Intelligence in order to respond to the needs of judges in the field of judicial knowledge.

PALAVRAS CLAVE: Inteligencia Artificial. Kendoj. Poder Judicial. España

KEYWORDS: Artificial intelligence. Kendoj. Judges. Justice. Spain

1 | IA en Justicia: ni sobrestimar ni subestimar

Cuando nos referimos al impacto de la Inteligencia Artificial (en adelante, IA) en los sistemas de justicia se sugiere, en este estado de cosas, tomar una posición prudentemente vigilante, esto es, una actitud ecléctica.

Esta actitud debe traducirse, en no sobrestimar, pero tampoco subestimar el irreversible impacto y penetración de la IA en el funcionamiento de la justicia.

Ni debemos sobreestimar la cuestión, como algo apocalíptico e irreversible para el pensamiento humano racional (tesis extrema), ni distópico (tesis futurista) ni revolucionaria (tesis posibilista) en cuanto a que comporta un radical cambio paradigma en cuanto a la transformación (o supresión) de métodos de pensar, porque hay razones para pensar que las máquinas no pueden pensar como nosotros lo hacemos, de ahí que se hable del “mito de la IA”¹. Pero tampoco podemos subestimarla en el sentido de concebirla como un simple cambio de métodos de trabajo para abreviar y simplificar la toma de decisiones.

La transformación digital constituye una prioridad para todas las Administraciones Públicas y, en especial, también para el Poder Judicial. Así se constató de manera temprana por el Consejo Consultivo de Jueces Europeos Informe N.º 14 (2011) “*Justicia y tecnologías de la información*”, y de manera específica se ha abordado en el último informe de 1 de diciembre de 2023 (Dictamen número 26) del CCJE titulado «*Aller de l'avant : l'utilisation de la technologie d'assistance dans le système judiciaire*»², en el que se afronta el uso de IA en los tribunales de Justicia. El título de este dictamen es toda una declaración de principios en su título: habla de “asistencia tecnológica”, por tanto, se trata de destacar la idea del control humano, y la tecnología como instrumento.

Dentro de este contexto, la reciente incorporación de tecnologías disruptivas, que permiten la incorporación de desarrollos y utilidades basadas en Inteligencia Artificial y *Machine Learning*, plantea retos y oportunidades que inciden directamente sobre el ejercicio de la función jurisdiccional.

Desde el Poder Judicial debemos reflexionar ante esta evolución que llega a constituir un auténtico cambio de paradigma. No solo se trata de dotar a la Carrera judicial de los medios más avanzados para que la justicia pueda incrementar su eficiencia, sino que es preciso reflexionar sobre el alcance que estos medios presentan y su posible repercusión sobre la propia

¹ Erik J. Larson. *El mito de la Inteligencia Artificial. Por qué las máquinas no pueden pensar como nosotros lo hacemos*. Shackleton -books-, 2022

² Silguero Estagnan, Joaquín. “*El impacto de la inteligencia Artificial desde la perspectiva del Poder Judicial*”. Cendoj, Consejo General del Poder Judicial. 2023

independencia judicial³.

La clave está en que los mecanismos de IA, su uso y resultados, pasen por el tamiz de su juridificación, de su acomodación a los principios esenciales de todo ordenamiento jurídico asentado en los estándares de un Estado de Derecho. A título ilustrativo. Pensemos, por ejemplo, la indudable trascendencia procesal y sobre los derechos de las partes, que pueden llegar a tener la incorporación o alteración de los tradicionales medios de prueba, la desinformación o la utilización de utilidades basadas en Inteligencia Artificial para la investigación de ilícitos penales, por citar tan solo algunos ejemplos.

2 | Tres temáticas y problemas de la ia en los sistemas de justicia

Si hubiera que exponer un catálogo de problemas sobre el impacto de la IA en los sistemas de justicia, cabría hacer la siguiente exposición, sin que su orden implique preferencia alguna.

Primera problemática: sobre la codificación normativa y rango normativo y de protección de la IA.

Aquí se suceden varios interrogantes: a) ¿Estamos ante un Derecho supranacional? La IA no es asunto de un país o de un grupo de países. Esto un tema global. También es un asunto de poder. Y de ahí el segundo interrogante: b) ¿Cuál debe ser el papel de las organizaciones internacionales? C) Una tercera cuestión, de técnica legislativa, se abre paso ¿se debe codificar normativa completa y detallada o solo principios a modo de estándares? ¿Derecho rígido o derecho dúctil o *soft law*? ¿Codificación de principios, pero también de reglas? ¿Son necesarias y oportunas leyes integrales sobre justicia digital? Una elocuente manifestación la hallamos en el hecho de la Unión Europea ha tomado recientemente la decisión de legislar sobre IA en su plenitud. Ahora bien, en caso de ser necesarias y oportunas dichas normas ¿cómo acompañar la elaboración técnica de estas leyes de justicia digital con las realidades organizativas -situación de los Tribunales- y la legislación procesal?

Segunda problemática: el campo de influencia de la IA modernización en los sistemas judiciales y el factor humano como *última ratio*.

Las tecnologías digitales no afectan solamente a la libertad, sino también a la igualdad y a la justicia⁴. El uso del algoritmo envuelve toda una

³ Mark Coeckelbergh, *La filosofía política y la inteligencia artificial. Una introducción*. Cátedra. 1ª edición, Madrid, 2023. Pág. 53.

teoría de filosofía política, si el sesgo del algoritmo creado no es justo. Pensemos en actuaciones automatizadas simples (o de puro trámite) y proactivas (más complejas, que pueden ser asistidas o no asistidas) en un proceso judicial. Los sistemas de justicia se construyen sobre los principios y reglas procesales. La incidencia de la IA será más o menos intensa, compleja y decisiva según en qué fases del proceso nos hallemos. Menos intensa en la fase de tramitación. Más compleja y complicada en la fase valorativa o de prueba. Muy intensa e invasiva si se emplea en la fase de decisión. Estos escenarios de aplicación de la IA tienen su predicamento y manifestaciones en la denominada justicia robótica (terminología más que discutible por suponer un oxímoron), tribunales *on line*, y la llamada justicia predictiva, con los límites que ello puede comportar.

Tercera problemática: IA aplicada a la Administración de Justicia en un Estado de Derecho: contenidos y límites legales y éticos: hacia un nuevo paradigma en la vertebración técnica del Derecho.

Esta tercera cuestión plantea a su vez tres escenarios de reflexión, que seguidamente apuntamos de modo esquemático:

(A) En primer término, los límites de la IA con relación a los derechos fundamentales, principalmente el derecho a la intimidad y la protección de datos. Para la construcción jurídica de esta materia brotan nuevos interrogantes ¿serían necesarios nuevos derechos? ¿son necesarias nuevas herramientas de protección? Lo primero parece que no, pues contamos con normativa y categorías jurídicas al respecto. Lo que hace falta no es tanto que se crean nuevas categorías jurídicas, cuanto que se apliquen las normas que recogen esos derechos fundamentales y exista una cultura jurídica de aplicación y respeto de las normas tanto en los operadores jurídicos como en la ciudadanía y, en particular, en los usuarios o quienes se ven involucrados en esa tecnología. A tal efecto, resulta importante tomar como marco de referencia la Carta Europea de Ética sobre el uso de la IA en sistemas judiciales y su entorno, adoptada por la Comisión Europea por la Eficiencia de la Justicia (CEPEJ, diciembre 2018). Tal y como considera dicho organismo, la aplicación de Inteligencia Artificial (AI) en el campo de la justicia puede contribuir a mejorar la eficiencia y la calidad y debe implementarse de manera responsable que cumpla con los derechos fundamentales garantizados, en particular, en la Convención Europea sobre Derechos humanos (CEDH) y el Convenio del Consejo de Europa sobre la protección de datos personales. Para el

CEPEJ, es esencial asegurar que AI siga siendo una herramienta al servicio del interés general y que su uso respete los derechos individuales. En cuanto a los principios fundamentales que deben respetarse en el campo de la AI y la justicia, se señalan los siguientes:

1. Respeto a los derechos humanos, incluyendo la no discriminación. Aspecto este de fundamental importancia a fin de preservar los derechos de las personas vulnerables. En suma, debe tratarse de soluciones que impulsen el respeto de los derechos garantizados por el Convenio Europeo de Derechos Humanos y el Convenio núm. 108 del Consejo de Europa.
2. Principio de calidad y seguridad. En el sentido de que debe ser posible procesar los datos mediante el aprendizaje automático sobre la base de originales certificados, garantizándose la integridad de estos datos en todas las etapas del procesamiento. Una buena práctica en este sentido es la creación de equipos multidisciplinares que impliquen a jueces y a investigadores en ciencias sociales e informática, tanto en la fase de redacción como de dirección y en la aplicación de las soluciones propuestas.
3. Principio de transparencia: dado que es preciso garantizar la accesibilidad y comprensión de las técnicas de procesamiento de datos, así como en la posibilidad de que las autoridades o expertos independientes realicen auditorías externas. Incluso se defiende el establecimiento de un sistema de certificación, que se renueve periódicamente.
4. Principio “bajo el control del usuario”: elemento éste que es preciso reforzar con el objeto de reforzar la posición del usuario. En este sentido, la situación deseable pasaría porque el usuario fuera informado, de manera clara y comprensible, de que las soluciones propuestas por las utilidades de Inteligencia Artificial son o no vinculantes.

Todos estos principios son los que, a juicio de la CEPEJ, deben garantizarse al procesar las decisiones judiciales y los datos. A lo que deberíamos añadir, que, sin ellos, difícilmente podría sustentarse una aplicación basada en Inteligencia Artificial que pueda ser respetuosa con los derechos humanos y la perspectiva de defensa de la persona.

(B) En segundo lugar, hay que identificar y, en su caso, evaluar los riesgos de la generalización de la IA aplicada a la Justicia con relación los usuarios/as de la justicia: la “brecha digital” con relación a personas vulnerables y “pies descalzos”.

- (C) En tercer lugar, los estados a la hora de introducir mecanismos de IA en los sistemas procesales deben tener clara la hoja de ruta. No deben legislar aluvionalmente, sin sistema ni coherencia. Ello comporta abordar otro nuevo interrogante ¿hasta qué punto se puede violentar la IA, como fenómenos proteicos (que cambia de formas o ideas) en sus múltiples y sibilinas variantes los tradicionales principios procesales de oralidad, intermediación, publicidad y principio dispositivo, por ejemplo, en el proceso civil?
- (D) Finalmente en la toma de decisiones, el uso de la IA se puede convertir en un instrumento obligado de utilización, como fuente de consulta, para tomar decisiones. Y, es más: su no empleo para la toma de decisiones puede derivar responsabilidades a quienes no la han utilizado. Pensemos, por ejemplo, no solo en el ámbito previo de la decisión judicial, sino en las decisiones de cogobernaba en el ámbito corporativo de unas sociedades mercantiles.

Las posibilidades de mejora de los actuales sistemas de reconocimiento de voz, la percepción visual, las funcionalidades de predicción de textos, así como otras tareas se ven impulsadas gracias a estas tecnologías disruptivas. Sin embargo, no debe olvidarse el fundamental principio de todo desarrollo basado en Inteligencia Artificial como es su aproximación basada en la defensa de la persona. Solo a través de aplicaciones que cumplan con esta premisa podrá alcanzarse una adecuada vertebración de la Inteligencia Artificial.

3 | Realidades concretas: experiencias en el poder judicial español (consejo general del poder judicial. Especial referencia al sistema *KENDOJ*

En el caso de España, el Centro de Documentación Judicial (CENDOJ)[5], órgano técnico dependiente del Consejo General del Poder Judicial de España, ha desarrollado varias utilidades basadas en Inteligencia Artificial con la finalidad de dar respuesta a las necesidades de los jueces en el ámbito del conocimiento judicial. También viene realizando dentro de su plan de formación de la carrera judicial cursos y seminarios sobre IA y ha participado en trabajos científicos sobre esta materia⁶.

Nos limitaremos a exponer el Sistema Kendoj sobre anonimización de

⁴ <https://www.poderjudicial.es>

⁵ *El derecho y la Inteligencia Artificial*. VV.AA. Alfonso Peralta, Leopoldo Salvador Torres y Francisco Herrera (coordinadores). Universidad de Granada, Editorial Universidad de Granada, Consejo General del Poder Judicial. Granada, 2022

⁶ <https://www.poderjudicial.es>

documentos e hipervinculación de legislación y sentencias. También está el uso de modelos de resoluciones judiciales frecuentes en los órdenes judiciales social, familia, penal, administrativo, y el de herramienta de cálculo de indemnizaciones laborales⁷.

La *suite* de utilidades denominada KENDOJ, “Knowledge Extractor for CENDOJ”, que se encuentra ya disponible en España para todas las personas que integran los Tribunales de justicia. Inicialmente desarrollada para Jueces y Magistrados, se benefician de la misma también otros colectivos en virtud de convenio (Letrados de la Administración de Justicia, Fiscales y Abogados del Estado).

Es una herramienta desarrollada con técnicas de inteligencia artificial, que permite, partiendo de un documento Word, (RTF) estructurarlo, detectar entidades, seudonimizarlo y validar dicha seudonimización, detectar vínculos a jurisprudencia y legislación y clasificarlo en base a voces del Tesauruso jurídico Cendoj. Veamos cada una de dichas funcionalidades por separado.

Por lo que se refiere a la funcionalidad de seudonimización, Kendoj permite aplicarla de manera automática a cada resolución judicial (actualmente disponible para formatos doc o rtf), para casos en que a un usuario/a le interesa guardar en sus archivos una resolución judicial sin los datos personales.

Esta herramienta permite plena autonomía al usuario y no interfiere en el proceso de tratamiento que el CENDOJ aplica a las decisiones judiciales que forman parte del Fondo Documental, que en la actualidad cuenta ya con 8,5 millones de decisiones judiciales.

En la práctica, se han detectado casos de uso entre los miembros de la carrera judicial, Letrados de la Administración de Justicia, fiscales y oficinas de comunicación (prensa) e Instituciones, tales como Ayuntamientos, Diputaciones... interesados en guardar las resoluciones judiciales en que son partes, seudonimizadas.

También presenta utilidad para el propio Consejo General del Poder Judicial dado que pueden darse supuestos en que sea interesante guardar una resolución judicial sin los datos procesales, evitando así entrar en colisión con la normativa vigente en materia de protección de datos.

Téngase en cuenta, para comprender lo que aquí decimos, que, en el Derecho español, el artículo 235 de la Ley Orgánica del Poder Judicial menciona que el acceso a las resoluciones judiciales por quienes no son parte se llevará a cabo “previas las medidas de protección de los datos personales”.

Es decir, que el acceso a las resoluciones judiciales, o a determinados extremos de las mismas, o a otras actuaciones procesales, por quienes no

son parte en el procedimiento y acrediten un interés legítimo y directo, podrá llevarse a cabo previa disociación, anonimización u otra medida de protección de los datos de carácter personal que las mismos contuvieren y con pleno respeto al derecho a la intimidad, a los derechos de las personas que requieran un especial deber de tutela o a la garantía del anonimato de las víctimas o perjudicados, cuando proceda.

En otros casos, es el propio Letrado de la Administración de Justicia quien hace uso de esta funcionalidad, dado que en base al artículo 235 bis LOPJ, le puede interesar anonimizar unos datos sí y otros no, y esta utilidad de “seudonimización” facilitaría la corrección de los datos ocultados y la emisión del certificado previsto en dicho artículo:

Así se establece en el artículo 235 bis 2 que, en los casos previstos en el apartado anterior, el Letrado de la Administración de Justicia emitirá certificado en el que se harán constar los siguientes datos:

- a) Los que permitan la identificación del proceso judicial.
- b) Nombre y apellidos o denominación social del condenado y, en su caso, del responsable civil.

El régimen normativo se completa con la previsión del artículo 236 quinquies LOPJ referida a quién debe adoptar las medidas necesarias para la supresión de los datos personales.

A tal efecto, se alude a los Jueces y Magistrados, los Fiscales y los Letrados de la Administración de Justicia, conforme a sus competencias, podrán adoptar las medidas que sean necesarias para la supresión de los datos personales de las resoluciones y de los documentos a los que puedan acceder las partes durante la tramitación del proceso siempre que no sean necesarios para garantizar el derecho a la tutela judicial efectiva, sin que, en ningún caso, pueda producirse indefensión.

Por último, también utilizan esta aplicación basada en Inteligencia Artificial, las Oficinas de comunicación. Debe tenerse en cuenta que éstas, en el ejercicio de sus funciones de comunicación institucional, deberán velar por el respeto del derecho fundamental a la protección de datos personales de aquellos que hubieran intervenido en el procedimiento de que se trate. Para cumplir con su finalidad, podrán recabar los datos necesarios de las autoridades competentes.

En cuanto a la funcionalidad relativa a la clasificación y citas estructurales de documentos, se realiza utilizando herramientas de Inteligencia Artificial, un enriquecimiento documental que ofrece la información de voces que identifican la materia a la que hace referencia y, así mismo, posibilita

el enlace con las citas a legislación y jurisprudencia a la base de datos del Fondo Documental.

KENDOJ es un sistema de procesamiento de documentos jurídicos en español, fundamentalmente resoluciones judiciales, cuyo objetivo es extraer datos normalizados que sirvan para enriquecer el Fondo Documental del CENDOJ.

Dicha aplicación ofrece la utilidad de clasificar un documento judicial asignándole voces del tesoro Cendoj de manera automática, facilitando al usuario una primera visión del documento por los términos jurídicos asignados según el “score” de relevancia, sin necesidad de una lectura íntegra y detallada.

Esta utilidad facilitará además la recuperación de la información de la legislación citada, con acceso al artículo y a la norma con un solo click, y la recuperación de las resoluciones judiciales citadas que de forma automática queden resueltas, con acceso a la legislación y jurisprudencia del fondo documental.

El acceso a la jurisprudencia del fondo documental le permitirá acceder a los valores añadidos del análisis y entre ellos a la línea de resoluciones del caso y por tanto recuperar información sobre una posible sentencia posterior a la citada; así mismo la vinculación de la norma citada a la legislación del fondo documental le permitirá el acceso a la información sobre su vigencia, y posibles modificaciones y versiones de esta.

También este caso, se han detectado casos de uso que demuestran la aceptación de esta utilidad. Así, cabe citar los miembros de la carrera judicial, Letrados de la Administración de Justicia, fiscales, el propio Consejo General del Poder Judicial, Instituciones, que podrán obtener un conocimiento rápido del contenido jurídico de un documento y un acceso directo a los vínculos de legislación y jurisprudencia citados en el mismo (querrela, demanda de cualquier jurisdicción, recursos...)

De conformidad con el artículo 236 sexies de la LOPJ, los sistemas automáticos de clasificación documental están incluidos dentro de las políticas de apoyo a la justicia. En este sentido se dispone que “el Ministerio de Justicia y las Comunidades Autónomas con competencias en la materia, dentro de las políticas de apoyo a la Administración de Justicia y desarrollo de la gestión electrónica de los procedimientos, podrán realizar el tratamiento de datos no personales para el ejercicio de sus competencias de gestión pública, incluyendo el desarrollo e implementación de sistemas automáticos de clasificación documental orientados a la tramitación procesal, con cum-

plimiento de la normativa de interoperabilidad, seguridad y protección de datos que resulte aplicable”.

En todo caso, lo cierto es que junto a los beneficios que la incorporación de estas nuevas utilidades supone, también surgen cuestiones legales y éticas a las que es preciso dar respuesta previamente a su implantación.

A través de las funcionalidades basadas en Inteligencia Artificial, se puede trabajar de forma más eficiente y automatizar tareas repetitivas. De este modo puede darse una respuesta más útil y eficiente a la situación de saturación que sufren los Tribunales.

Ahora bien, existen determinados retos a los que es preciso dar cumplida respuesta. Por ejemplo, los algoritmos utilizados no siempre resultan transparentes, lo que dificulta -o imposibilita- su supervisión. Otro riesgo evidente es que los sistemas basados en Inteligencia Artificial permitan perpetuar los sesgos y la discriminación.

De ahí que resulte fundamental controlar la calidad de los datos suministrados. Es por ello por lo que conforme al Reglamento sobre Inteligencia Artificial que actualmente se está negociando en el seno de la Unión Europea, el uso de Inteligencia Artificial en el sistema de justicia se considera una aplicación de alto riesgo.

Un límite resulta claro y es que los sistemas de Inteligencia Artificial no deben reemplazar el componente humano. Es más, me atrevería a señalar que es difícil defender la credibilidad del propio sistema, si no se mantiene dicho factor. Los jueces en el futuro no solo necesitarán dictar sentencias, sino que además deberán decidir en qué medida la Inteligencia Artificial es utilizada en su trabajo.

En la actualidad no existen todavía estándares en cuanto al uso de la Inteligencia Artificial en el ámbito judicial. Han comenzado esos ensayos en el marco de la XXI Cumbre judicial Iberoamericana (Perú, 2023)[8], en el que se aprobaron en la Asamblea Plenaria las “aplicaciones de IA en los Poderes judiciales”, un Código de Ética de IA” y respuestas al cuestionario “Lago de datos”⁹. Asimismo, la Comisión de Coordinación y Seguimiento de la Cumbre Judicial Iberoamericana en su reunión de 17-19 de noviembre de 2023, Santo Domingo, República Dominicana)[10] ha asumido la encomienda de formular una propuesta en el ámbito del a IA y aplicación a la Administración de Justicia. Tampoco cabe obviar que las aplicaciones de Inteligencia Artificial son con frecuencia desarrolladas de forma comercial

⁷ <https://www.cumbrejudicial.org/>

⁸ <https://www.cumbrejudicial.org/node/696>

⁹ *Post scriptum* a la intervención realizada en Colóquio “Tribunais e Inteligência Artificial – Uma Odisseia no século XXI”, integrado no Ciclo de Colóquios do Supremo Tribunal de Justiça, realizou-se no dia 11 de maio, Lisboa

y no como tecnología de código abierto. Todas estas cuestiones deben ser abordadas antes de la incorporación de las tecnologías basadas en inteligencia artificial en el trabajo de los Tribunales de justicia, sin olvidar el factor humano. •





LA APLICACIÓN ÉTICA A LA INTELIGENCIA ARTIFICIAL EN LAS RELACIONES JURIDICAS



Luís Maria Benitez Riera
Ministro da Corte Suprema, Paraguai
Miembro de la Comisión de Coordinación Y Seguimiento
de la Cumbre Judicial Iberoamericana

Clique para assistir.



Los Estados miembros de la UNESCO adoptan el primer acuerdo mundial sobre la Ética de la IA. **En noviembre de 2021, los 193 Estados miembros de la Conferencia General de la UNESCO** adoptaron la Recomendación sobre la Ética de la Inteligencia Artificial, el primer instrumento normativo mundial sobre el tema.

Este Acuerdo, sienta la discusión en los ejes transcendentales de análisis previo, a la adopción de sistemas que ofrezcan una alternativa de aplicación de la Inteligencia artificial a la resolución de conflictos judiciales.

No obstante, RECOMIENDA que los Estados Miembros apliquen, de manera voluntaria, las disposiciones de la presente Recomendación mediante la adopción de las medidas adecuadas, en particular las **medidas legislativas** o de otra índole que puedan ser necesarias, de acuerdo con la práctica constitucional y las estructuras de gobierno de cada Estado, con el fin de dar efecto en sus respectivas jurisdicciones a los principios y normas enunciados en la Recomendación, de conformidad con el derecho internacional, incluido el derecho internacional de los derechos humanos;

Esta recomendación nos plantea la necesidad de manera previa de la efectiva protección de los datos personales, a través de una Ley integral, que regule aspectos inherentes al ser humano; y al mismo tiempo reglar nuevas formas o herramientas de investigación que incorporen a la inteligencia artificial a la producción de pruebas, que luego deban ser admitidas en juicios.

Algunos argumentan que la IA podría ayudar a crear un sistema judicial penal más justo, en el que las máquinas pudieran evaluar y sopesar los factores relevantes mejor que los humanos, aprovechando su velocidad y la ingesta de grandes cantidades de datos. Por lo tanto, AI tomaría decisiones basadas en decisiones informadas sin prejuicios ni subjetividad.

Puede ocurrir que, en algunas ocasiones, los seres humanos decidan depender de los sistemas de Inteligencia Artificial por razones de eficacia, pero la decisión de ceder el control en contextos limitados seguirá recayendo en los seres humanos, ya que estos pueden recurrir a los sistemas de IA en la adopción de decisiones y en la ejecución de tareas, pero un sistema de IA nunca podrá reemplazar la responsabilidad final de los seres humanos y su obligación de rendir cuentas. Por regla general, las decisiones de vida o muerte no deberían cederse a los sistemas de IA

Pero hay muchos desafíos éticos:

- Falta de transparencia de las herramientas de IA: las decisiones de IA no siempre son inteligibles para los humanos.
- La IA no es neutral: las decisiones basadas en la IA son susceptibles de imprecisiones, resultados discriminatorios, sesgos incrustados o insertados.
- Prácticas de vigilancia para la recopilación de datos y la privacidad de los usuarios de los tribunales.
- Nuevas preocupaciones por la equidad y el riesgo de los Derechos Humanos y otros valores fundamentales.

Una vez identificados, los principales desafíos, que, desde el campo ético, presenta la Inteligencia Artificial, debemos ver como esta herramienta de vincula con los **operadores** de la justicia, tiene el potencial de optimizar los tiempos de la **justicia**, y con los **usuarios** permite que el abogado ahorre tiempo, evite inconsistencias en sus documentos, que aplique un mayor rigor jurídico y se reduce el riesgo de tomar decisiones incorrectas, los niveles de accesibilidad y al mismo tiempo la sostenibilidad de estos sistemas.

En la actualidad, existen sistemas cognitivos que analizan datos y jurisprudencia, constituyendo un apoyo fundamental para la labor de los jueces. Lo que ahora se discute es si un algoritmo puede actuar de forma autónoma como juez.

La labor de juzgar comporta, al menos, **cuatro actividades**: seleccionar el material jurídico relevante aplicable, interpretarlo y aplicarlo al caso, determinar los hechos (pruebas) y las relaciones causales y calificar

judicialmente los hechos probados.

Hay capacidades esenciales del hombre de las que carecen las máquinas: moral; razonamiento por analogía; empatía; compasión; ambigüedad del lenguaje.

Si bien, el derecho, desde su campo, debe adaptarse a la implementación de esta nueva tecnología, resulta controversial permitirle una activa participación en toma de decisiones, sin siquiera entender cómo lo hace. Así, uno de los principales retos del derecho es definir hasta dónde y cómo permitirse el uso de esta inteligencia y sus límites frente a los derechos fundamentales.

En nuestra opinión, cualquier proyecto de exploración sobre el empleo de la inteligencia artificial y otras tecnologías asociadas en la administración de justicia precisa, en primer lugar, la adopción de un enfoque basado en una apreciación realista y pragmática de las capacidades y limitaciones de la inteligencia artificial en su estado actual de desarrollo, dejando de lado tanto posiciones utópicas o excesivamente optimistas sobre sus potencialidades futuras como aquellas que rechazan de principio o se muestran totalmente escépticas acerca de las posibilidades de su empleo.

En segundo lugar, un enfoque holístico, que comprenda una visión global del sistema judicial y de las distintas funciones y tareas que implica el desarrollo del proceso judicial

Una pregunta que debe responderse, como lógica de los análisis previamente realizados en las as decisiones judiciales han de ser motivadas, lo cual significa que deben estar fundadas en argumentos jurídicos o razones jurídicamente aceptables. Pero **¿puede la inteligencia artificial proporcionar ese tipo de razones para justificar sus decisiones?** el problema de la justificación es fácilmente resoluble, con las capacidades que a día de hoy ofrecen el aprendizaje automático y el procesamiento del lenguaje natural.

Parece insostenible, desde una perspectiva mínimamente garantista, que se pueda aceptar como un fundamento jurídicamente válido para la adopción de una medida tan grave como la privación (aun provisional) de libertad una predicción automatizada sobre el comportamiento futuro, para lo cual será una exigencia que las herramientas de Inteligencia Artificial cumplan rigurosamente con el principio de legalidad del ordenamiento jurídico garantizando una verdadera protección a la dignidad, la libertad, el debido proceso, la presunción de inocencia, entre otros pilares fundamentales del estado de derecho.

Si bien es cierto, que en los países de América del sur hemos avanzado de manera diversa en la aplicación de la Inteligencia Artificial para la resolución de conflictos jurídicos, destacamos la necesidad de continuar el debate ético, que nos permita limitar estrictamente las acciones y por otro lado las herramientas que deban ser desarrolladas, minimizando los sesgos y las asimetrías. •



PODER JUDICIAL DE PUERTO RICO: DOS INICIATIVAS BASADAS EN INTELIGENCIA ARTIFICIAL



Sigfrido Steidel Figueroa¹
 Director Administrativo de los Tribunales de Puerto Rico
 y Juez del Tribunal de Apelaciones
 Miembro de la Comisión de Coordinación Y Seguimiento
 de la Cumbre Judicial Iberoamericana

Clique para assistir.



La progresiva incorporación de tecnología informática en el quehacer judicial es una constante hace varias décadas en prácticamente todos los poderes judiciales. La pandemia causada por el COVID-19 aceleró ese proceso por la necesidad de garantizar el distanciamiento físico y por la gran cantidad de recursos económicos destinados a encarar la emergencia. En poco tiempo se reformularon procedimientos, normas administrativas y reglas procesales para adaptarlas a un escenario nuevo en el que las premisas de la actividad judicial que parecían inmutables perdieron vigencia².

La inmediatez presencial de los actores del escenario judicial ha perdido espacio frente a la actividad virtual. El papel ha dejado de ser el soporte material principal para comunicar ideas y argumentos entre las partes y los jueces. Y la autenticidad y confiabilidad de un dictamen, orden o resolución no dependen de un sello adherido a un papel. No extraña que a consecuencia de los cambios que ha supuesto la incorporación de tecnología informática en la litigación la judicatura haya enfrentado controversias jurídicas novedosas, no solo relacionadas a las medidas preventivas tomadas en el contexto del COVID-19, sino también en temas sustantivos, procesales y probatorios.

La inteligencia artificial es en tiempos recientes la nueva frontera que miramos con curiosidad y escepticismo. Presenta oportunidades, retos y,

¹ Las expresiones vertidas en este escrito no representan la posición oficial del Poder Judicial de Puerto Rico, ni la de sus jueces, empleados y funcionarios.

² Véase, Cumbre Judicial Iberoamericana, *El COVID19 y la administración de la justicia Iberoamericana*. Disponible en: www.cumbrejudicial.org (última visita, 18 de diciembre de 2023).

según el criterio de diversas voces, algunos peligros. Indistintamente, el tema se aborda desde múltiples ámbitos como lo son las dimensiones ética, jurídica, sociológica, administrativa y tecnológica, entre otras. Aunque el debate acrecienta y el tema más escabroso parece ser la posibilidad de que la actividad judicial sea asumida en todo o en parte por máquinas, no existe duda de que en múltiples aspectos es una herramienta que permite acelerar y facilitar procesos de todo género, incluidas tareas relacionadas al proceso de juzgar.

El Poder Judicial de Puerto Rico, como muchos otros poderes judiciales, se ha interesado en las oportunidades que la inteligencia artificial ofrece para agilizar procesos. A tono con ello, se desarrolló una aplicación que utiliza herramientas basadas en inteligencia artificial para asistir a la ciudadanía en la preparación de peticiones de órdenes de protección en casos de violencia de género. Mediante un avatar que simula a un funcionario judicial y las conversaciones por chatbot los usuarios aclaran dudas sobre el proceso. Al completar el formulario inteligente, se genera de forma automática la petición de la orden judicial de protección, y con ello, el expediente electrónico del caso. La aplicación se complementa con la sala judicial virtual — establecida en Puerto Rico durante la pandemia causada por el COVID-19 —, la cual facilita que la persona que promueve la orden de protección pueda, desde cualquier lugar de Puerto Rico y mediante el uso de una computadora o unidad móvil inteligente, participar en una audiencia judicial de manera remota, e incluso fuera del horario laboral regular, en la que un juez evaluará los méritos de su petición. Como se aprecia, se trata de una herramienta que agiliza la presentación de los documentos necesarios para obtener una orden de protección en asuntos en los que es necesario actuar con rapidez en consideración a los riesgos potenciales para la persona que solicita la intervención judicial.

Un segundo proyecto en desarrollo que prevemos que estará disponible en el 2024, integrará en una misma base de datos las decisiones emitidas por los tribunales de los tres niveles jerárquicos de Puerto Rico. Con apoyo de inteligencia artificial, la herramienta identificará de manera automática, desde que se presenta electrónicamente un nuevo caso o asunto, sentencias y resoluciones emitidas en casos con controversias judiciales similares. Tras identificarlas, seleccionará las más afines o relevantes al nuevo caso y las pondrá a disposición del juzgador para su examen y consideración en su proceso decisorio. La herramienta no sustituye el criterio judicial, solo le asiste en el proceso investigativo.

A diferencia de motores de investigación comerciales que solo integran decisiones de los foros apelativos, esta herramienta contará con una base de datos más amplia, pues integrará decisiones de tribunales de todos los niveles jerárquicos, incluido el Tribunal de Primera Instancia en donde se presenta la mayor variedad de asuntos y controversias judiciales. Asimismo, contrario a los motores de investigación comerciales, no requerirá que el usuario incorpore criterios de búsqueda. La propia inteligencia artificial identificará los conceptos relevantes del nuevo caso presentado y automáticamente identificará las sentencias y resoluciones emitidas en casos similares por cualquier juez, sea del Tribunal de Primera Instancia, del Tribunal de Apelaciones o del Tribunal Supremo. De este modo, la falta de destrezas investigativas del usuario no impedirá al juzgador contar con las decisiones más relevantes para decidir un caso. Claro está, corresponderá al juzgador confirmar la similitud fáctica y jurídica, cuáles de estas decisiones son precedentes obligatorios que no puede soslayar, cuáles han sido revocadas por foros de mayor jerarquía, así como los cambios normativos que hayan afectado la vigencia de la decisión identificada por la herramienta. En un futuro no muy lejano la propia inteligencia artificial también facilitará estos procesos.

Estos esfuerzos van de la mano de otros, como la continua mejora de las plataformas digitales de manejo del expediente electrónico, la adquisición de equipo adecuado a las nuevas herramientas y necesidades, y la educación continua del personal judicial y de la comunidad jurídica en general. Somos conscientes de que estas aplicaciones en desarrollo no agotan las posibilidades de la inteligencia artificial. No obstante, son un paso importante en nuestro país a la luz de nuestras posibilidades y necesidades. •





TRIBUNAIS E INTELIGÊNCIA ARTIFICIAL



Gabriela Cunha Rodrigues

Juíza Desembargadora

Chefe do Gabinete do Presidente do Supremo Tribunal de Justiça

Clique para assistir.



A visão do digital na área da Justiça vai oscilando entre o revolucionário e o conservador, o entusiasmo e o ceticismo.

Richard Susskind aponta para uma ideia de transumanismo, em que humanidade, tecnologia, robótica, nanotecnologia e genética convergiram.

Mais céticos, Antoine Garapon e Jean Lassègue publicaram um livro (*Justice Digitale*) com a inquietação espelhada na cinta que o envolve: «Aceitariam ser julgados por algoritmos?»

A inteligência que comanda a análise probatória e a decisão não dispensa a consciência do juiz.

No entanto, a inteligência artificial poderá ser um instrumento importante em etapas decisivas para a formação da decisão.

Na tramitação e consulta de dados, o uso de ferramentas de inteligência artificial potencia resultados positivamente superiores aos atingidos por mão humana.

Significa poupança de trabalho burocrático e liberta o juiz para a sua verdadeira função de julgar.

Mais humano, menos máquina...

O Supremo Tribunal de Justiça de Portugal tem procurado ser pioneiro na implementação de técnicas de inteligência artificial.

Apresentou uma candidatura em 2020 ao Projeto IRIS/SAMA – Informação, Racionalização, Integração e Sumarização, desenvolvido pelo INESC-ID – Instituto de Engenharia de Sistemas e Computadores e financiado através do Programa Operacional COMPETE 2020, enquadrado no

Programa Portugal 2020, e pelo Fundo Social Europeu.

Está em fase final de produção e de testagem.

O âmbito do Projeto IRIS/SAMA, atualmente definido, é o desenvolvimento de ferramentas que cumpram três grandes objetivos:

- **Apoio à análise de processos e elaboração de acórdãos;**
- **Apoio na elaboração de sumários, na sugestão de descritores, na anonimização dos acórdãos e na transcrição para formato digital de acórdãos antigos que só existam em papel;**
- **Divulgação pública da jurisprudência do Supremo Tribunal de Justiça através da sua publicação no *site* deste Tribunal, na forma de documentos previamente anonimizados.**

Neste contexto, foi efetuada investigação sobre o estado da arte das tecnologias relevantes nos domínios da ciência dos dados e da inteligência artificial.

Definiu-se o desenvolvimento de cinco serviços, a suportar por quatro aplicações:

- **Serviços de «biblioteca digital»**, com o objetivo de disponibilizar a pesquisa e o acesso aos textos em formato digital e aos metadados das decisões proferidas pelo Supremo Tribunal e Justiça
 - **Serviço de divulgação ao público em geral** dos textos anonimizados dessas decisões, permitindo a sua descoberta por mecanismos variados de pesquisa e navegação em índices;
 - **Serviço de pesquisa restrito de Jurisprudência do Supremo Tribunal de Justiça**, apenas acessível na rede interna do Tribunal, em que se permite a descoberta e o acesso aos textos das decisões não anonimizadas.
- **Serviço de apoio à anonimização de acórdãos**
- **Serviço de apoio à sumarização e geração de descritores de acórdãos:** serviço que sugere a um utilizador que lhe forneça um texto e indica quais os parágrafos que podem ser considerados mais relevantes para uma tarefa de sumarização desse texto.

Este serviço é realizado através de três módulos interligados de forma transparente para o utilizador, designadas de «Sumarizador», «Descritores» e «Interface».

- **Serviço de apoio à realização de OCR** (acrónimo de «Optical Character Recognition»): serviço que permite a um utilizador que lhe

forneça um conjunto de imagens de apoio na tarefa de transcrição para um formato de texto editável.

Todas as soluções foram desenvolvidas utilizando código livre e, sempre que possível, código aberto.

Todas elas fazem uso de «tecnologia web», o que implica que são destinadas a ser instaladas em servidores acessíveis em redes com tecnologia normalizada da *Internet* e a ser acedidas através de «navegadores de *www*» usuais.

Já a elaboração de sumários em linguagem formal, com base no texto original do acórdão, é uma tarefa extremamente complexa, representando o desafio de maior exigência científica deste projeto.

Um sumário é uma construção intelectual única, obviamente relacionada com o texto do acórdão, mas bastante distinta no seu conteúdo.

Há que reconhecer que este desafio vai para além do estado da arte das tecnologias atuais, pelo que deveria ser reconhecidamente classificado como de investigação científica avançada.

Todas as aplicações desenvolvidas podem vir a ter um potencial de utilização mais alargado, tornando-as eventualmente úteis para outros tribunais em Portugal e noutros países, designadamente os países e territórios de língua portuguesa.

Atualização dos critérios de seleção e pseudonimização das decisões judiciais

O Supremo Tribunal de Justiça publica tendencialmente todos os acórdãos proferidos desde finais de 2020.

Tem promovido vários encontros e reuniões com o Conselho Superior da Magistratura e todos os Encarregados da Proteção de Dados da área da Justiça no sentido de superar as dúvidas que a aplicação prática dos critérios de pseudonimização (expressão mais apropriada do que anonimização) suscita.

Tem beneficiado da experiência da Divisão de Documentação e Informação Jurídica do Supremo Tribunal de Justiça na aplicação do *software* «anonymizer», em estreita colaboração com a *Data Juris*.

Fruto da enriquecedora experiência prática do Supremo Tribunal de Justiça, o Plenário do Conselho Superior da Magistratura, por deliberação de 11.4.2023, procedeu recentemente à atualização das regras essenciais da pseudonimização das decisões judiciais, de acordo com os critérios de orien-

tação adotados pelo Supremo Tribunal de Justiça.

É inegável e evidente a importância da publicação das decisões judiciais, essencial para a transparência, o conhecimento e o escrutínio da atividade dos tribunais.

Potencia a promoção de uma cultura jurídica baseada na jurisprudência, tornando possível que as decisões dos tribunais sejam objeto de estudo pelos profissionais do Direito, pelas Universidades e pela doutrina.

O seu tratamento automatizado permite uma multiplicidade de funcionalidades pela análise de dados que a publicação *online* de forma estruturada permite.

Nestes procedimentos automatizados, onde fica a tomada de decisão humana?

A pseudonimização pode ser efetuada com recurso a inteligência artificial, mas pela experiência já longa do Supremo Tribunal de Justiça, concluo que, em qualquer caso, haverá sempre lugar à revisão humana antes da publicação.

Nesta Odisseia no Século XXI, termino com uma citação do poeta espanhol António Machado:

“Caminante, no hay camino. Se hace camino al andar!... •



ES NECESARIA LA ELABORACIÓN DE PAUTAS PARA EL DESARROLLO DE LOS SISTEMAS DE INTELIGENCIA ARTIFICIAL EN MATERIA JUDICIAL?



John Perez Brignani

Ministro da Corte Suprema do Uruguai
Miembro de la Comisión de Coordinación Y Seguimiento
de la Cumbre Judicial Iberoamericana

Clique para assistir.



En materia de inteligencia artificial Uruguay se encuentra en la tapa de definición de proyectos y bases para la implementación de la inteligencia artificial a los distintos proyectos que emerjan en un futuro próximo

En ese marco se celebró un convenio con la UBA a fin de realizar un programa que permita aquilatar la peligrosidad de los denunciados en casos de violencia doméstica. Esta herramienta constituirá un aporte importante para el juez encargado de aplicación de las medidas cautelares necesarias para salvaguardar a la víctima

En efecto se desarrollara una prueba piloto de un sistema de inteligencia artificial que, tomando en cuenta los distintos perfiles y datos de víctimas y victimarios, permita establecer el posible grado de peligrosidad concreto en cada caso. Ello permitirá clasificar por el grado de gravedad, las denuncias por violencia contra las mujeres y las personas de su grupo familiar y, así, reducir tiempos y optimizar los tiempos de análisis y optimizar la clasificación en función de variables relevantes para tomar decisiones (que en muchos casos se basan en evaluar o reevaluar el riesgo y luego dictar las medidas correspondientes) y tomar las medidas adecuadas en cada caso concreto.

El proyecto comprende un módulo basado en técnicas de aprendizaje automático y/o visión artificial y/o el uso de expresiones regulares algorítmicas a los fines de realizar: Detección y clasificación automatizada sobre datos o patrones de información relevantes de las denuncias escaneadas (colaborar

con la automatización para evaluar la situación de la víctima: por ejemplo, denuncia en comisaría de policía) y perfil del denunciante.

Para el diseño, incubación, desarrollo y despliegue de la prueba piloto, se trabajará en conjunto con dos poderes judiciales (juzgados) y/u oficinas de administración pública de dos países de la región como pilotos, con el fin de generar recomendaciones y lecciones aprendidas que luego permitan servir de base para adaptar, desarrollar más robustamente, escalar y, eventualmente, integrar el sistema en otros poderes judiciales y/u oficinas de administración pública de estos dos países, o de otros países de la región de América Latina y el Caribe. Los países considerados inicialmente son Argentina y Uruguay. El proyecto se basa en documentar la experiencia y disponibilidad el código con el que se trabaje, más allá de que se priorizará el uso de herramientas *open source*.

Ahora bien a mi modesto entender más importante que el desarrollo de sistemas en materia de inteligencia es la elaboración de pautas para su posible concreción y desarrollo. En efecto si bien la utilización de la inteligencia artificial trae muchos beneficios conlleva que se puedan ver afectados diversos derechos fundamentales tales como reforzar y reproducir los prejuicios ya existentes, vulnerar los derechos de los ciudadanos a un proceso justo entre otros.

Ejemplo de estas últimas son por ejemplo los programas de predicción de actividad delictiva en los cuales puede llegar a un alto grado de discriminación racial que conlleva a acentuar aún más los perfiles discriminatorios.

La capacidad de las máquinas para simular la capacidad de entender, aprender, razonar y actuar para resolver situaciones específicas o interactuar con procesos y personas, puede originar los conflictos antes mencionados y otros debido a que las máquinas son programadas por seres humanos y la selección que los mismos realizan.

En este orden juega un importante papel la ética, que en el caso de inteligencia artificial podemos definirla como un subcampo de la ética aplicada que estudia los problemas éticos que plantea el desarrollo, despliegue y utilización de la IA. Su preocupación fundamental es identificar de qué modo puede la IA mejorar o despertar inquietudes para la vida de las personas, ya sea en términos de calidad de vida o de la autonomía y la libertad humana necesarias para una sociedad democrática. (Cfm ISSN 2444-8478 • <http://doi.org/10.12795/IESTSCIENTIA.2021.i01.19> Derechos humanos, ética y transparencia algorítmica, Evelyn Téllez Carvajal 372 (Grupo in-

dependiente de Expertos de Alto Nivel sobre Inteligencia Artificial, 2018, p. 13 IUS ET SCIENTIA • 2021 Vol. 7 • Nº 1 • pp. 370 – 386).

En el contexto de la inteligencia artificial, el respeto de la dignidad humana implica que todas las personas han de ser tratadas con el debido respeto que merecen como sujetos morales, y no como simples objetos que se pueden filtrar, ordenar, puntuar, dirigir, condicionar o manipular. En consecuencia, los sistemas de IA deben desarrollarse de un modo que respete, proteja y esté al servicio de la dignidad física y mental de los seres humanos, el sentimiento de identidad personal y cultural y la satisfacción de sus necesidades esenciales. **(Grupo independiente de Expertos de Alto Nivel sobre Inteligencia Artificial, 2018, p. 15).**

Corolario de ello es que en el caso de la inteligencia artificial el derecho de cualquier individuo a conocer en detalle los criterios mecánicos que le son aplicados para decidir si posee o no un rasgo determinado, para decidir si cumple o no con un criterio predicado, es esencial. El problema se suscita entre el derecho de los afectados y la posible confidencialidad y reserva de las partes decisorias. Debemos tener presente que por un lado, existe el derecho a la información, derecho humano universal, pero también existe la posibilidad de que esta información sea reservada. Ahora bien, ello no implica que las decisiones en las cuales se basa el sistema no deban explicarse, ya que se vería altamente vulnerado el derecho a la información y el principio de transparencia, elementos claves para una adecuada defensa de los intereses en juego. No obstante cabe destacar que “El grado de necesidad de explicabilidad depende en gran medida del contexto y la gravedad de las consecuencias derivadas de un resultado erróneo o inadecuado”. **(Grupo independiente de Expertos de Alto Nivel sobre Inteligencia Artificial, 2018, p. 16).**

Otro principio directamente implicado es el principio de equidad en el desarrollo, despliegue y utilización que conlleva a una distribución justa e igualitaria de los costes y beneficios, pero sobre todo que determinadas personas no se vean sometidas a un proceso de estigmatización o discriminación injustificadas

Respecto de este último aspecto opera en el caso de la inteligencia artificial el principio de diversidad y de no discriminación en virtud del cual “Para hacer realidad la IA fiable, es preciso garantizar la inclusión y la diversidad a lo largo de todo el ciclo de vida de los sistemas de inteligencia artificial. Además de tener en cuenta a todos los afectados y garantizar su participación en todo el proceso, también es necesario garantizar la igual-

dad de acceso mediante procesos de diseño inclusivos, sin olvidar la igualdad de trato. Este requisito está estrechamente relacionado con el principio de equidad”. (**Grupo independiente de Expertos de Alto Nivel sobre Inteligencia Artificial, 2018, p. 23**).

“La diversidad y la inclusión desempeñan un papel esencial al desarrollar los sistemas de IA que se utilizarán en el mundo real. Es crucial que, a medida que los sistemas de IA vayan desempeñando una mayor cantidad de tareas por sí mismos, los equipos encargados de la adquisición o del diseño, desarrollo, ensayo, mantenimiento y despliegue de estos sistemas reflejen la diversidad de los usuarios y de la sociedad en general. Esto contribuye a garantizar la objetividad y la toma en consideración de las diferentes perspectivas, necesidades y objetivos. Lo ideal es que la diversidad no sólo se materialice en los equipos en términos de género, cultura y edad, sino también de antecedentes profesionales y conjuntos de competencias”.

Si bien las decisiones tomadas en algoritmo debe ser los más neutrales y basarse en datos objetivos, no es menos cierto que su utilización implica una serie de elecciones y supuestos que no son neutrales ya que se basan en la selección de determinados valores precisos y no otros. Es por ello que la transparencia juega un rol esencial a fin de determinar la posible existencia de discriminación o que se base en datos erróneos de la realidad.

En primer lugar, el principio de cognoscibilidad, según el cual toda persona tiene derecho a conocer la existencia de procesos automatizados de toma de decisiones que le conciernen y, en su caso, a recibir información significativa sobre la lógica utilizando el principio de comprensibilidad, es decir, la posibilidad, de retomando la expresión del Reglamento, Europeo recibir “información significativa sobre la lógica utilizada”.

La decisión automatizada “produce efectos jurídicos que no afectan o afectan significativamente a una persona”, ésta tiene derecho a que esta decisión no se base únicamente en este proceso automatizado (artículo 22 del Reg.). En este sentido, en todo caso debe existir en el proceso de toma de decisiones una contribución humana capaz de controlar, validar o denegar la decisión automática.

Ahora bien en el caso de mi país y de muchos países latinoamericanos carecen de una regulación adecuada de la aplicación de la Inteligencia Artificial. Es por ello que creo que deberíamos comenzar a trabajar en tal sentido antes de emprender la diagramación e instrumentación de programas en tal sentido.

A tales efectos considero que podría tomarse como base, sin perjuicio de

las particularidades de cada país el reglamento europeo que regula de la Inteligencia Artificial.

Considero que los diversos países, que integran la Cumbre, deberíamos propugnar tomar como base el citado reglamento porque considero que el mismo contempla las distintas hipótesis y bajo su égida se encuentran tres importantes países de la cumbre como son Portugal, España y Andorra.

En efecto el mismo tiene cinco principios claves para la regulación a saber:

1. Principio del respeto de los derechos fundamentales
2. Principio de no discriminación
3. Principio de calidad y seguridad
4. Principio de transparencia
5. Principio de garantía de la intervención humana

Debemos tener presente que los sistemas de IA deben estar destinados a ayudar a una autoridad judicial en la investigación e interpretación de hechos y de la ley, así como en la aplicación de la ley a un conjunto concreto de hechos.

Pero su aplicación no puede implicar en grado alguno la disminución de derechos sino todo lo contrario . Así como los justiciables tienen derecho a un proceso eficaz y eficiente , para lo cual la inteligencia artificial puede realizar numerosos aportes , no es menos cierto que tales procesos impliquen la vulneración de los derechos de los mismos.

No debemos olvidar que el proceso es un instrumento para la actuación de los derechos La lógica y las estadísticas ayudan, pero resultan inadecuadas con respecto a la historia humana subyacente. Esto requiere la evaluación y el control de un ser humano. Es una cuestión de planes y perspectivas si se quiere: así cómo sólo la máquina puede recibir y reprocesar exhaustivamente la información adquirida de otra máquina, pero solo un ser humano puede comprender a otro ser humano. •





ENCERRAMENTO



Agostinho Torres

Juiz Conselheiro do Supremo Tribunal de Justiça

Clique para assistir.



Exº Sr Presidente do STJ

Excelências - Ilustres Convidados- Minhas senhoras e meus Senhores

*

Ainda embalados pelo som de uma excelente música que nos transportou ao expoente dos sentidos (e humano, por enquanto), conclui-se assim, em fim de tarde, este relevantíssimo ciclo de colóquios subordinado ao tema TRIBUNAIS E INTELIGÊNCIA ARTIFICIAL – UMA ODISSEIA NO SÉCULO XXI

1. Como bem recordarão, foi o bioquímico e escritor de ficção científica russo-americano Isaac Asimov (1920-1992) quem popularizou, em meados do século XX, o conceito de inteligência artificial. Aqui e agora, desde logo, o termo Odisseia, no título deste colóquio, sugere-nos de novo o **deslumbramento** que sempre suscita uma viagem para um novo desconhecido e, se tal acontece também com novos sentimentos de esperança para a sobrevivência da humanidade, ao mesmo tempo desencadeia uma profunda preocupação com a capacidade humana de controle da sua própria criação.

2. Na verdade, é consabido que todos os avanços científicos, em particular no campo das novas tecnologias, podem trazer benefícios ou malefícios, consoante a for-

ma como são utilizados, o que **desperta duplamente esperanças e receios**. Veja-se o que terá sido desde logo a descoberta de como dominar o fogo, utilizável tanto para aquecer como para ferir ou destruir. (**Na biologia** podem criar-se organismos que nunca existiram, ou novos vírus capazes de trazer graves riscos à saúde de seres vivos ou permitem também criar novos medicamentos capazes de reduzir o sofrimento ou levar à cura. **Na química** podem apoiar a criação de novos materiais, que podem tanto corroer ou destruir objetos, como causar danos que provocam acidentes que podem lesar ou matar seres vivos, mas podem também criar próteses que reduzem o desgaste físico e que melhoram a vida das pessoas. **Na área jurídica**, novas leis ou novas interpretações de leis existentes podem pôr em risco o meio ambiente, a liberdade e o bem-estar das pessoas. **Pesquisas na área de comunicação** podem aumentar a disseminação e a assimilação de notícias falsas...etc, etc, etc.).

3. Ora, em face daquele deslumbramento, terá razão a sabedoria chinesa quando afirma ser “uma maldição vivermos em tempos interessantes”? Na verdade, o temor das capacidades da IA tem crescido em intensidade tal que já levou a reações em série de governos, empresas e organismos sociais. Veja-se, por exemplo, na Europa, onde a Itália bloqueou o funcionamento do ChatGPT, (aplicativo criado pela OpenAI) ou a circulação muito recentemente de uma petição on line (da Fair Trials) a exigir que a nova Lei de IA a discutir no PE bloqueie e proíba o uso pelas autoridades de sistemas de predição de perfis criminais de indivíduos.

4. O certo é que, por mais que investigadores brilhantes como Wozniak, criador do primeiro computador da Apple, prefiram que se interrompa o progresso da IA, a história também ensina que impor barreiras à inovação pode ser um caso perdido. Na Imprensa, o jornal i (português) publicou uma edição produzida pelo ChatGPT. Na rádio, a Futuri Media criou o software RadioGPT e lançou uma rádio que gere e anima com vozes criadas por IA (ou clonadas a partir de vozes de animadores e jornalistas). E assume-se como melhor do que qualquer outra rádio, porque a sua IA sabe tudo sobre todos os artistas. Para quem ouve, se não souber que as vozes são artificiais, vai achar que são mesmo animadores e jornalistas que estão do outro lado do microfone. Na TV, a primeira jornalista pivô criada e gerada por IA acabou de iniciar funções na China no People’s Daily (controlado pelo Estado chinês) e está a fazer emissão em contínuo, 24 horas por dia. No cinema, as startups Deepcake e Respeecher criam clones digitais (imagem e áudio) de atores e celebridades. O ator Bruce Willis clonou a sua imagem para participar num filme publicitário e o ator James Earl Jones clonou a sua voz para continuar a dar voz a Darth Vader

na saga StarWars. Ou seja, a IA vai garantir a “imortalidade”... e não está longe o dia em que vamos escolher os atores que queremos ver num filme ou série. Na música, o Google, que está há anos a trabalhar em inteligência artificial e de onde se esperam muitas novidades em breve, já anunciou que vai lançar a MusicLM, que criará músicas com base nos nossos pedidos.

Embora difícil a tarefa, importa pois parar um pouco e tentar compreender o potencial disruptivo de uma tecnologia e/ou dimensionar os riscos que ela representa.

5. Os «**Sistemas de inteligência artificial**» são concebidos para funcionar com um cada vez maior nível de **autonomia**, (e que, com base em dados e entradas fornecidos por uma máquina e/ou pelo homem, deduz como atingir um determinado conjunto de objetivos usando aprendizagem de máquina e/ou abordagens baseadas em lógica e conhecimento e produz saídas geradas pelo sistema, como conteúdo (sistemas de IA generativos), previsões, recomendações ou decisões, influenciando os ambientes com os quais o sistema de IA interage.)” mas não são poucos os **perigos a eles associados**. Entre os mais marcantes estão a concentração de poder nas mãos de poucas empresas, o desaparecimento de empregos pela automação de atividades, a maioria será alterada, (*apesar de que muitos empregos novos também serão criados*), a disseminação descontrolada de ataques cibernéticos ou o desenvolvimento de armas autónomas ou mesmo, a crer nas recentes notícias, as investigações sobre a criação de textos por leitura de pensamentos. ¹

6. Perante esta fulgurante realidade dos novos avanços da IA, a Comissão Europeia vem incentivando os Estados-Membros a modernizar seus sistemas de ensino e formação e a apoiar as transições no mercado de trabalho com base no Pilar Europeu dos Direitos Sociais (PEDS) para lidar com questões relacionadas com o impacto da Inteligência Artificial (IA) no futuro do trabalho e na legislação. Também o Grupo Europeu de Ética em Ciência e Novas Tecnologias (EGE) e a Agência UE dos Dtos Fundamentais estão trabalhando juntos para proporem recomendações ou directivas à Comissão Europeia sobre o futuro desenvolvimento de políticas relacionadas com a IA e com questões éticas, legais e sociais. Por sua vez, formou-se a AI Alliance para orientar o trabalho do Grupo de Especialistas de Alto Nível em Inteligência Artificial (AI HLEG) procurando-se estabelecer um diálogo político aberto sobre esta tecnologia.

¹ <https://pt.euronews.com>

“(...) novo sistema de IA que pode traduzir a actividade cerebral de alguém num fluxo contínuo de texto. “Para um método não invasivo, este é um verdadeiro salto em frente em comparação com o que foi feito antes, que é tipicamente uma única palavra ou frases curtas”, disse Alex Huth, professor assistente de neurociência e ciência da computação na UT Austin, e um dos autores do artigo.”

7. Perante o surgir de novas questões éticas e legais neste panorama, ligadas à responsabilidade ou à tomada de **decisões potencialmente tendenciosas**, manipuladoras ou baseadas em dados falsos, procurou-se com este colóquio examinar os **dilemas envolventes e o enquadramento legal** que resultam do espantoso desenvolvimento da IA. Foram partilhadas estratégias para determinar o equilíbrio entre a inovação em IA e a garantia de proteção dos utilizadores. Reflectiu-se também sobre **impacto no judiciário**, nas suas vantagens e vulnerabilidades bem como sobre a **possível substituição dos juízes por máquinas** dotadas de IA.

8. Com o objetivo de identificar as oportunidades, vulnerabilidades e desafios associados ao impacto da implementação dessas tecnologias na profissão de juiz, no processo judicial e no ordenamento jurídico, foi realizada uma pesquisa com 360 juízes e apresentados os resultados inéditos do inquérito, efectuando-se ainda um percurso pelos argumentos que sustentam ou se opõem a uma teoria da substituição jurisdicional, partilhando com o público uma reflexão crítica.

II

9. Mas que novo mundo é este o da IA?

-“*Não acho seja o fim do mundo, mas é um desafio novo, que se tem vindo a complicar de dia para dia*”, referiu recentemente o comandante Santos Fernandes, chefe de divisão de Planeamento do Estado Maior da Armada, na conferência “Confiar no jornalismo, fugir à desinformação”, que decorreu na sede da Lusa, em Lisboa. Defendendo que temos sempre tendência para correr atrás do prejuízo, lembrando os escândalos com algoritmos e redes sociais e apesar de acreditar que existe já uma “*consciencialização para o problema*”, alertou que este é “***um novo mundo mas também uma enorme ameaça***”.

10. Sobre esta matéria, **a IA já está presente nas nossas vidas** (*através de uma Alexa, Siri ou mesmo de um Joaquim*) através dos assistentes virtuais disponíveis nos telemóveis, tablets e outros ‘gadgets’. **Já se fala em direitos humanos para os ‘robots’** para que tenham “um pouco mais de controlo” sobre o seu funcionamento, (ficção científica?), o que **acarreta problemas de ética** e vem *alterar a ideia de que a inteligência artificial é apenas uma máquina*.

11. Também o recurso a sistemas de (IA) e de tomada de decisão automatizada (ADM) é cada vez mais frequente por parte de autoridades europeias no campo da aplicação da lei e da justiça criminal, nomeadamente para traçar o perfil das pessoas, prever seu suposto comportamento futuro e avaliar seu suposto “risco” de

criminalidade ou reincidência no futuro.

(Discute-se o uso de sistemas de policiamento preditivo e seu impacto prejudicial. Esses sistemas podem ser usados para prever o risco em áreas como policiamento e justiça criminal, mas essas previsões podem levar a vigilância constante, interceptações frequentes de cidadãos e buscas, sanções (multas), interrogatórios, detenções, processamento criminal e prolação de sentenças. Eles também podem influenciar a decisão e controle de liberdade condicional ou mesmo levar a punições não criminais, como a negação de assistência social ou outros serviços essenciais e até mesmo a remoção de crianças de suas famílias. Organizações de direitos humanos como a “Fair Trials” pediram já a proibição do uso de IA e ADM (*Automatic Decision Making*) pelas autoridades de justiça criminal exigindo **salvaguardas legais rigorosas para o uso das mesmas** pois, segundo elas, os **sistemas preditivos** usados no policiamento e na justiça criminal **minam direitos fundamentais e reforçam a discriminação**. Os dados usados para criar, treinar e operar sistemas de policiamento preditivo **refletem preconceitos sistêmicos, institucionais e sociais**. (*vide* o caso *Delia*, um sistema preditivo na Itália, usando dados de etnia para traçar o perfil e “prever” a criminalidade futura.)

12. A IA traz novos desafios. E porquê?

Porque permite que as máquinas “aprendam” e tomem ou implementem decisões sem intervenção humana, funcionalidade esta que se tornará **padrão** em muitos tipos de bens e serviços, desde smartphones a carros automatizados, robôs ou aplicativos online. No entanto, **há o perigo de as decisões tomadas por algoritmos sem intervenção humana poderem resultar de dados incompletos** e, portanto, **não confiáveis**, de serem adulteradas por ciberataques, serem tendenciosas ou simplesmente equivocadas. **A aplicação irrefletida da tecnologia** à medida que ela se desenvolve pode conduzir, portanto, a resultados problemáticos e à relutância dos cidadãos quanto ao respectivo uso.

13. Contudo, a inteligência artificial pode contribuir para vastos benefícios econômicos e sociais em todo o espectro de indústrias e de atividades sociais. O seu uso pode fornecer vantagens competitivas importantes para empresas e apoiar resultados social e ambientalmente benéficos, por exemplo, **na saúde, agricultura, educação e aprendizagem, gestão de infraestruturas, energia, transporte e logística, serviços públicos, segurança, justiça, eficiência energética e de recursos e mitigação e adaptação às alterações climáticas**.

14. Mas, como “não há bela sem senão”, ao mesmo tempo, e dependendo das cir-

cunståncias relativas à sua aplicação e utilização específicas, a IA **pode gerar riscos** e prejudicar interesses e direitos protegidos pelo direito (dano material e/ou imaterial). A recente **estratêgia europeia de IA** em matéria de regulação pretende deixar claro que a **confiança** surge como um pré-requisito para garantir uma abordagem centrada no ser humano: “a IA não é um fim em si mesma, mas uma ferramenta que deve servir as pessoas” com o objetivo final de aumentar o bem-estar humano. Para conseguir isso, **a confiabilidade da IA deve ser garantida através da integração dos valores europeus fundamentais** (*pluralismo, a não discriminação, a tolerância, a justiça, a solidariedade e a igualdade, o respeito pela dignidade humana, liberdade, democracia, igualdade, Estado de direito e respeito pelos direitos humanos, incluindo os direitos das pessoas pertencentes a minorias*) nos quais as nossas sociedades se baseiam, **na forma como a IA se desenvolve.**

15. Daí que à **UE caberá garantir uma forte estrutura regulatória** que defina o padrão global para a IA, **centrada pois no ser humano**, sendo premente um quadro jurídico da União com regras harmonizadas nesta matéria. Os sistemas de IA podem ser facilmente implantados em vários setores da economia e da sociedade, inclusive transfronteiriços, e circular por toda a União. Nessa medida e em antevisão, alguns Estados-Membros já exploraram a adoção de regras nacionais para garantirem que a IA seja segura, desenvolvida e utilizada em conformidade com as obrigações em matéria de direitos fundamentais.

16. Neste plano estratégico, a **Comissão Europeia adotou uma proposta de regulamento** que visa estabelecer regras harmonizadas em matéria de inteligência artificial (Lei da Inteligência Artificial, AIA) em 21 de abril de 2021. Por isso que, já em 25 de novembro de 2022 - o Conselho da UE aprovou uma **versão de compromisso da proposta** de Lei de Inteligência Artificial. Embora detectando a existência de divergências na definição dos sistemas de IA e dificuldades na definição do conceito de autonomia destes, o Conselho acredita que a definição não deve incluir certos tipos de software existentes.

17. Assim,

Em sede de Práticas proibidas de IA - o texto da proposta de Lei de Inteligência Artificial agora considera como tais **o uso de IA para pontuação social** de atores privados de sistemas de IA que exploram as vulnerabilidades de um grupo específico de pessoas, incluindo pessoas vulneráveis devido à sua situação social ou econômica. Porém, **já quanto à proibição do uso de sistemas de identificação biométrica** remota ‘em tempo real’ em espaços acessíveis ao público pelas

autoridades policiais **o texto da Lei de Inteligência Artificial proposta ainda refere na presente data que tal uso é estritamente necessário** para fins de aplicação da lei e para o qual as autoridades policiais devem ser excepcionalmente autorizadas a usar tais sistemas.

18. Noutro plano e mais além Atlântico, a figura que mais críticas tem dirigido à evolução tecnológica associada à IA tem sido, no entanto, quem mais tem contribuído para a sua aplicação. (*Observador 23 abr. 2022*). Trata-se de Elon Musk. Este, em setembro de 2020 dizia que “tanto os smartwatches, como os smartphones podem ser considerados como as tecnologias de ontem, ao contrário da ligação neural desenvolvida pela Neuralink, Inc., que almeja conectar o cérebro humano directamente aos dispositivos electrónicos e que, como tal, pode ser considerada uma das tecnologias do futuro.”

19. Este alegado **novo produto do grupo Tesla** ainda que, para já, do ponto de vista tecnológico, possa parecer irrealista, levanta inúmeras e **preocupantes questões**. Desde logo, destaca-se a da eventual conexão com a **Starlink, Inc.**, que visa permitir ao utilizador **navegar na internet sem depender de uma ligação com as tradicionais redes de telecomunicações**. Ou, ainda, **a conexão neural que se pretende estabelecer entre os smartphones e os chips cerebrais desenvolvidos pela já referida Neuralink, Inc.**, (empresa que, recentemente, foi acusada de sujeitar macacos a tratamentos e experiências considerados cruéis), é aquela que mais questões éticas, morais ou sociais levanta e a que mais controvérsia gera.

20. A longo prazo, o **objectivo passa por utilizar a tecnologia de interface neural para fundir a consciência humana com inteligência artificial e, com isso, alcançar uma espécie de simbiose**.

21. Mas os eventuais **benefícios do desenvolvimento** da IA ultrapassam as eventuais **desvantagens até que ponto?** Sendo a maioria de computadores em princípio, “pirateável”, os avanços tecnológicos não são em si dissuasores e acabam por constituir um atractivo ou mesmo um desafio. Veja-se **por exemplo** o recente aumento de ciberataques a organizações ou crescente prática de cibercrimes e as respectivas graves consequências em termos de danos financeiros e reputacionais.

22. Ora, para um futuro não muito longínquo, já se pensa pois em **chips implantados no cérebro humano fundidos com inteligência artificial**. As conse-

quências serão inimagináveis. Como será possível garantir a protecção dos nossos dados pessoais e industriais em caso de ciberataque? Conjugando a teoria freudiana do sistema de percepção-consciência, até que ponto será possível garantir a protecção do próprio “processo psíquico consciente”, inclusive do nosso estado prévio inconsciente? (*vide* filme *Inception*, - *Leonardo DiCaprio*- 2010?)

23. Hoje em dia, cada pormenor das nossas vidas já é monitorizado e analisado pelos nossos próprios dispositivos electrónicos, com vista à recolha e armazenamento de dados para efeitos de criação de perfis psicográficos detalhados, **mas o nosso cérebro contém informações que, por natureza, não são ainda passíveis de ser registadas. A conexão directa entre o nosso cérebro, o nosso dispositivo electrónico e a acção da inteligência artificial, pode vir a estabelecer uma espécie de via de acesso à nossa consciência** e, com isso, além de poder vir a permitir o registo dessas informações, **pode ainda vir a permitir que hackers assumam a porta de entrada e mesmo o controlo da nossa mente**, o que inclui a memória e o processo de tomada de decisão e execução.

24. Imaginem-se as **consequências da aplicação de tal tecnologia no âmbito de determinado grupo de pessoas**, a título de exemplo, em termos de eleições políticas e sentido de voto, ou mesmo em termos de adesão ou apoio a outros movimentos políticos mais ou menos pacíficos, mais ou menos activistas e, mesmo, mais ou menos agressivos e violentos. Ou mesmo até as ilimitadas possibilidades em termos militares e, ainda, em termos de vigilância policial e controlo social. Isso, para não falar já da eventualidade de, num cenário limite e apocalíptico, no futuro, a própria **inteligência artificial poder vir a assumir o controlo da interface neural e, assim, da nossa consciência...**

25. Não deixa, pois, de ser irónica a postura de Elon Musk perante as críticas que tem dirigido aos recentes desenvolvimentos tecnológicos e, ao mesmo tempo, ser dos que mais tem contribuído para o seu desenvolvimento e aplicação. Se, perante o perigo que a inteligência artificial representa, já em 2018, advertiu que ela pode vir a tornar-se num “**«ditador imortal»** do qual a humanidade «nunca irá conseguir escapar»” e em 2020, de novo alertou dizendo a «**inteligência artificial irá ultrapassar os humanos em menos de cinco anos**», **não obstante continuou** a desenvolver, sem limitações conhecidas, esta tecnologia, e a ampliar as suas eventuais formas de aplicação. (ao nível da robótica impulsionado pela Tesla, Inc., em concreto o projecto Tesla Optimus e a criação de robôs humanóides).

26. Foi assim que, nesta curta linha de tempo os receios assinalados levaram já a Alta comissária da ONU para os Direitos Humanos a alertar a 15 de Set de 2021 para o facto de **alguns sistemas** que utilizam a Inteligência Artificial (IA) serem ou poderem ser **discriminatórios ou ameaçadores** de algumas liberdades, pedindo uma **moratória ao seu desenvolvimento** às empresas e aos Estados que desenvolvem a tecnologia de IA que aumentem a transparência dessa investigação.

27. Entretanto, o mundo “acordou” para esta nova realidade com o **ChatGPT, (Transformador generativo pré-treinado)** lançado em novembro do ano passado, não obstante as imagens criadas pelos programas Dall-E e Mid Journey também já terem gerado enorme entusiasmo na comunidade tech. Como exemplo revelador da utilidade e interesse global nas aplicações da inteligência artificial o **ChatGPT²** (recurso este que simula a linguagem humana) **cresceu de tal modo que atingiu um milhão de utilizadores em apenas cinco dias e 100 milhões em dois meses** (o Instagram demorou cerca de dois meses a atingir um milhão). Desde que foi lançado, o ChatGPT conquistou corações e mentes ao responder a estímulos escritos dos utilizadores **como se fosse uma pessoa real**.

28. Nas últimas semanas, tem sido avassalador o número de notícias sobre os avanços tecnológicos no sector da inteligência artificial (IA), em concreto sobre o lançamento do [GPT-4](#) (sem esquecer, claro, o lançamento da [Bing AI](#) ou, mesmo, do [Google Bard](#)). O crescimento tão rápido deste tipo de sistemas de IA generativa promete, assim, revolucionar a maneira como a sociedade interage com a tecnologia, ao mesmo tempo que gera grandes receios porquanto **é uma tecnologia muito diferente do que se construiu até hoje** e em muito pouco tempo vai, entre outras consequências económicas e sociais, substituir muitos postos de trabalho. *“Importante será manter “o conhecimento, cultura e criatividade como fundamentais para nos mantermos relevantes e nos diferenciarmos”, mas também o será “saber utilizar as ferramentas de IA para aumentar a produtividade e a inovação nos produtos e serviços que criamos.”* (na opinião de Nuno Ribeiro, managing partner da Instinct a 28 de Abril de 2023 na ver on-line Meios & publicidade).

29. A revolução da IA será, pois, talvez, **“a maior revolução vivida pela Humanidade**, com enorme impacto em todas as indústrias, nos comportamentos sociais e individuais. Enquanto uns se entusiasmam e outros levantam a voz contra o súbito desenvolvimento da IA (v.g. o G7 e recentes declarações sobre necessidade de um travão e moratória à investigação), já **Gil Shwed**, considerado o pai da

² O GPT Chat é um recurso de inteligência artificial que simula a linguagem humana, criado por um laboratório de pesquisas em inteligência artificial dos Estados Unidos, a OpenAI. O nome Chat GPT é uma sigla para “Generative Pre-Trained Transformer”.

firewall moderna, que filtra tráfego de redes e computadores na internet, em entrevista recente ao Expresso em 30 de Abril de 2023, defende que não é possível parar o desenvolvimento da inteligência artificial: “(...) **Não devemos desacelerar a inteligência artificial, imagine que há 30 anos alguém dizia para suspender a Internet?**”

30. Enquanto assim, uns com receio e outros mais entusiastas, permanecem as **interrogações**:

A criatividade original do “inventor humano” irá ficar limitada à programação deste tipo de sistema pela mera introdução de dados/algoritmos e/ou depender exclusivamente da qualidade e capacidade de processamento do sistema em causa, ou seja, dos recursos económicos do proprietário do “inventor artificial”?

31. Estas e outras questões ficam por responder mas, em tempos de incerteza, mais vale termos o bom senso de parar o suficiente para pensar se queremos que a máquina seja apenas a nossa criação ou se vamos permitir que, como Adão e Eva, ela se rebelde contra o seu Deus criador.

32. **Não estamos pois, já em época de mudanças mas na mudança de uma época** (*citando, embora noutro contexto, uma expressão do Papa Francisco*). Mas, lembrando aqui as palavras de T. S. Elliot, “*oxalá não deixemos de continuar a explorar mas, no termo da nossa exploração, que consigamos sempre chegar ao ponto de partida e conhecer esse lugar pela primeira vez.*”

33. Indo já longa esta Odisseia de deslumbramentos e entropias mas a propósito, sempre, de invenções, ousarei terminar esta apresentação com uma frase do Sr. Cons.º Laborinho Lúcio em recente entrevista à radio em que, a dado passo, sublinhou que: “*A maior invenção da humanidade não são as coisas mas as pessoas*”.

34. Por fim, deixa-se aqui um profundo agradecimento a todos os que partilharam o seu saber nas apresentações feitas ao longo do dia, Srs Profs **Dr. Paulo Novais, Arlindo Oliveira, Vera Lúcia Raposo, e Dr. Nuno Lima da Luz**, aos Exmos membros da **Comissão de Coordenação e Acompanhamento da Cimeira Judicial Ibero-Americana**, aos Exmos. Srs. **Conselheiros moderadores** dos painéis **Dr. Cura Mariano e Dra. Maria Olinda Garcia** e a toda a equipe que pensou, desenvolveu e executou este colóquio, em especial à **Dra. Juíza Sandra dos Reis Luís** que desenhou o programa, ao **Sr. Conselheiro João Miguel e Dra. Andreia Martinho** que estiveram na base da elaboração e implementação

do questionário e estudo lançados.

É também devido, sempre, um enorme agradecimento aos excelentes músicos Desidério Lázaro e Mário Delgado pelo prazer que nos proporcionaram e a todos os que, desse lado, nos acompanharam neste colóquio, presentes ou on line, e nos deram a honra e o prazer de nos fazerem sentir que valeu a pena esta viagem sobre o mundo tão deslumbrante quanto temerário da Inteligência Artificial.

Votos de um bom resto de tarde para todos. •







À MARGEM DO COLÓQUIO

IMPULSO Y CONSTITUCIONALIZACIÓN DE LOS NUEVOS DERECHOS FUNDAMENTALES



John Perez Brignani
Ministro da Corte Suprema do Uruguai

Clique para assistir.



Los continuos cambios jurídicos y tecnológicos que operan en la sociedad conllevan a que se plantee en la órbita jurisdiccional, si determinados derechos, originados como consecuencia de los mismos, gozan o no de protección constitucional.

Ello es así por cuanto el catálogo de derechos constitucionalmente protegidos no es concebido como una lista cerrada. En efecto dicho catálogo no sólo puede sino que debe ser ampliado por vía interpretativa cuando los avances jurídicos y tecnológicos así lo requieran para una adecuada protección de los derechos fundamentales del individuo ..

Debemos tener presente que en el derecho constitucional la actividad del intérprete juega un rol crucial, ya que los textos de la constitución se refieren en términos genéricos y condicionados al contexto histórico en el que fueron aprobados. Y sin una labor interpretativa de carácter extensivo muchos de esas situaciones, en tanto derechos fundamentales, quedarían sin protección.

Los derechos fundamentales constituyen categorías abiertas y permeables a la existencia de nuevos valores y derechos. Esto significa que deben evolucionar acompañando las nuevas realidades económicas, sociales, políticas, culturales y tecnológicas de los pueblos cuando las exigencias de respeto a la dignidad humana lo requieran conforme a las mismas.

No debemos perder de vista que el constitucionalismo contemporáneo se caracteriza por una dimensión ética, sustancial, rica, articulada y hete-

rogénea (derechos, principios objetivos y valores), y que los derechos fundamentales son una construcción compleja que tiende a generar posteriores derechos funcionales a la protección del interés de partida.

La interpretación constitucional asume pues, en estos casos, no sólo una centralidad, que no se da en otras disciplinas, sino que tiene una característica distintivamente creativa, sustancial e inevitable.

Otro aspecto importante a destacar es que el modelo de constitucionalismo de derechos es un modelo supranacional, transnacional y global, donde a través de una copiosa producción de pactos y convenciones nacionales y regionales, se regulan los derechos fundamentales o derechos humanos que constituyen el centro de la problemática del derecho internacional y que se integran necesariamente al ordenamiento interno.

Y en esa labor el Poder judicial asume un rol central y activo en la protección de los derechos fundamentales y en general en la adecuación del ordenamiento jurídico a los principios constitucionales. Ello por cuanto para que una norma sea considerada como atributiva de derechos debe ser interpretada.

El Poder judicial no es más concebido como un neutral y fiel aplicador de la ley sino más bien como un guardián de la constitución y de los derechos fundamentales, que se involucra en la esfera de su actuación sobre el ordenamiento en el que le toca actuar, buscando siempre la posibilidad de individualizar por vía interpretativa nuevos derechos que protejan a los individuos: No debemos perder de vista que el Poder Judicial en tanto garante de los derechos de los ciudadanos debe tener como eje central de su accionar el individuo y en especial la salvaguarda de sus derechos fundamentales.

La fidelidad del juez a la ley se encuentra sujeta a la condición de que el legislador sea a su vez respetuoso de la constitución y si ese respeto no acontece debe aplicar esta última.

Lo que el Poder Judicial debe asegurar es la más amplia tutela posible de derechos individuales y en ese orden el catálogo de derechos reconocidos es muy amplio y heterogéneo, susceptible de posteriores ampliaciones por vía de integración supranacional.

Debemos tener presente que el constitucionalismo contemporáneo acentúa la voluntad política constituyente que establece preceptos que no sólo aseguran y garantizan la protección de los derechos fundamentales, sino también preceptos que obligan a los órganos del Estado y los poderes públicos a promover los derechos fundamentales.

En este orden cabe recalcar que las constituciones en general, como la

de mi país contienen una cláusula en blanco que puede ser interpretada en el sentido de autorizar la individualizaciones de nuevos derechos no expresamente comprendidos en el catálogo constitucional más congruentemente con el respeto a los valores fundamentales del ordenamiento.

Ahora bien, más allá de aquellos derechos que son considerados como fundamentales desde el punto de vista formal al ser atribuidos por una norma superior por su estructura jurídica cabe preguntarnos ¿qué parámetros debemos tener en consideración para saber si como consecuencia de hechos no previstos expresamente nos hallamos o no ante un derecho fundamental susceptible de protección?

El primer aspecto que hay que tener presente es que no es necesario que un derecho esté configurado expresamente en la Constitución formal para ser un derecho esencial, humano o fundamental. En tal sentido, la interpretación integradora del sistema de derechos permite desentrañar el contenido de las carencias normativas, acudiendo a los valores, a los principios y a la razón histórica de la Constitución y su sistema de derechos. En efecto constituciones en general reconocen derechos anteriores y preexistentes a la existencia de la norma jurídica, ya que los derechos son inherentes a la naturaleza humana y a la dignidad de la persona humana. Asimismo, los derechos pueden no coincidir necesaria y literalmente con el elenco que surge del texto constitucional ya que pueden emerger de una fuente diversa como son los tratados internacionales.

Es importante recalcar que la expresión derechos humanos reenvía a la idea de que son derechos que pertenecen al ser humano por su condición de tal independientemente de cualquier circunstancia y reconocimiento y por consiguiente debemos analizar expresamente si el derecho sometido a nuestra jurisdicción encuadra dentro de tal hipótesis.

El constitucionalismo moderno debe tender a que exista una expansión de los derechos constitucionalmente protegidos y que los mismos converjan en una óptima convivencia e integración para una completa protección de la dignidad humana.

Por tanto, a efectos de una aplicación o en todo caso de un uso de una ley fundamental es necesario pasar de la incorporación de la ley a una regla de principio o a su concreción en un conjunto de reglas, o incluso en un principio menos indeterminado y genérico que el de partida. De aquí surge inmediatamente una pregunta crucial: quién tiene la competencia para llevar a cabo estas operaciones cronológicas? ¿Quién debe identificar los métodos de implementación que hacen legalmente posible el ejercicio de un derecho

fundamental en la práctica, y coordinarlo con otros derechos?

En primer lugar, parece evidente que la concreción de los derechos fundamentales, y la identificación detallada de su disciplina, es tarea primordial del legislador. Y esto sobre todo si en la disciplina de un derecho se consideran perfiles directamente relacionados con el uso de los recursos públicos.

La cultura jurídica del constitucionalismo contemporáneo, sin embargo, no se conforma con esto, debido a la supremacía axiológica de la constitución, y su naturaleza de fuente de derecho a todos los efectos, en este contexto cultural, el *interpositio legislatoris* no es considerada ni necesaria ni suficiente con respecto a la individualización de la disciplina de los derechos fundamentales. No es necesario, pues se considera del todo legítimo que la actividad de concreción de derechos fundamentales se realice en sede interpretativa cada vez que falte una precisa disciplina legislativa necesaria para el ejercicio del derecho, de hecho, para el constitucionalismo contemporáneo. la constitución es fuente del derecho a todos los efectos directamente aplicables en los tribunales, y negar la posibilidad de que, en ausencia de una ley, los derechos sean confirmados en los tribunales significaría permitir que la normatividad de la constitución sea neutralizada por la mera inacción del legislador **(Cfm Pino Giorgio Il costituzionalismo dei diritti Il mulino itirinatti 2017 pag130)**.

Por consiguiente en el caso de que de la norma surjan explícitamente los derechos no existiría ningún tipo de inconvenientes. El problema se plantea en el caso de que el derecho no surja claramente consagrado en la norma en forma expresa lo que conlleve a determinar si de su interpretación emerge o no implícitamente de su texto tal consagración.

Para ello una posibilidad de interpretación para determinar si nos hallamos o no ante un derecho fundamental es a partir del contenido del derecho, objeto de análisis considerando como fundamentales aquellos derechos cuyo contenido o fundamento axiológico encuadra dentro de los parámetros antes reseñados.

Una segunda posibilidad de definición de los derechos fundamentales consiste en argumentar su existencia por analogía a partir de algún derecho ya existente y a través de esa vía poder individualizar ulteriores derechos que no estén expresamente formulados por analogía.

Otra posibilidad consiste en el proceder de un derecho de partida expresamente formulado y partiendo de un material normativo existente individualizar un derecho fundamental particularmente importante de manera igual a otros derechos civiles y políticos.

Ahora bien la proclamación de un derecho no equivale a su realización y menos a su garantía y en esto las Cortes constitucionales juegan un papel clave ya que ante el reconocimiento de un derecho fundamental como tal, éste debe ser asegurado mediante la tutela jurisdiccional necesaria

En este orden una de las mayores dificultades a la cual debemos enfrentarnos como intérpretes es la de hallar un punto de equilibrio en la interpretación de la situación subjetiva y sus límites y sobre todo de los diversos derechos en juego en determinados hechos.

Y uno de los avances en los que se registran mayores desafíos para la creación y protección de derechos individuales es el tecnológico, ya que las innovaciones tecnológicas determinan la necesidad de la protección de los derechos humanos en el orden individual y colectivo frente al Estado y los privados.

Ello es así por cuanto si bien se desarrollan nuevos espacios para ejercer determinados derechos humanos como ser la libertad de expresión, acceso a la información entre otros, y la creación de nuevos derechos como el derecho de alfabetización digital, como forma de poder ejercer estos derechos, estos medios han incidido negativamente por ejemplo en el derecho a la privacidad y en la discriminación. En efecto el surgimiento constante de nuevas tecnologías que permiten un mejor ejercicio de los derechos fundamentales tiene como contracara que se vean afectados otros derechos como en el de la privacidad entre otros.

La incorporación de las nuevas tecnologías, que permiten la circulación de millones de datos en cuestión de segundos, conlleva la aparición de nuevas posibilidades de intromisión y control en la vida privada de los individuos, exigen la necesidad de reforzar los sistemas de protección y de seguridad electrónica que posibiliten el pleno y libre ejercicio de los derechos que protegen constitucionalmente la esfera privada de las personas, en especial los relativos a la intimidad, la inviolabilidad del domicilio y el secreto de las comunicaciones.

Ello implica que el legislador habrá de buscar un adecuado equilibrio entre la protección de tales derechos y otros derechos e intereses no menos relevantes, como son los derechos a una comunicación libre, el interés creciente que tiene la sociedad en la circulación de la información sobre los individuos y el también legítimo interés del Estado en garantizar la aplicación de la ley. Indudablemente que la exigencia de que se proteja la intimidad experimentará un fuerte aumento al hacerse realidad el potencial de las nuevas tecnologías, capaces de conseguir (incluso más allá de las fronteras

nacionales) información detallada sobre individuos de fuentes diversas, y de manipular dicha información. Sin la seguridad legal, la falta de confianza por parte de los consumidores socavarán sin duda el rápido desarrollo de la sociedad de la información.

Pero no sólo el legislador opera exclusivamente en ese campo sino que el Poder Judicial como tal tiene una importante labor en los casos sometidos a su consideración, al tener que determinar si un determinado derecho debe o no ser protegido y sus límites.

Otro importante problema en el cual operan ambos poderes es en el evitar la exclusión social mediante la falta de acceso a las nuevas tecnologías de determinados grupos sociales.

Esto plantea la necesidad de establecer salvaguardas jurídicas y de garantizar la cohesión de la nueva sociedad. Deberá garantizarse a todos un acceso equitativo a la infraestructura, al igual que la prestación de un servicio universal, cuya definición debe ir de la mano de la evolución tecnológica. La educación, la formación y la promoción desempeñarán necesariamente un rol central en estos aspectos.

La libertad informática en estos tiempos adquiere la categoría de nuevo derecho fundamental que tiene como propósito garantizar la facultad de los individuos para conocer y acceder a las informaciones que les conciernen archivados en bancos de datos, controlar su calidad, lo que implica la posibilidad de autodeterminación informativa, de corregir o cancelar los datos inexactos o indebidamente procesados y disponer sobre su transmisión. El control de la información personal está relacionado con el concepto de autonomía individual para decidir cuándo y qué información referida a una persona, puede ser objeto de procesamiento automatizado.

La expansión del concepto de ciudadanía digital implica no sólo el libre acceso y uso de la información, sino una interacción simple y completa con las Administraciones Públicas. Es por ello que la lucha contra la exclusión digital debe ocupar un lugar esencial, mediante no sólo la educación sino también propugnando el acceso a medios tecnológicos.

Esto a su vez ha traído como consecuencia la necesidad de la fijación de nuevos límites en cuya tarea los operadores judiciales jugamos un importante papel.

Y en este orden uno de los mayores escenarios donde se plantean estos tipos de colisión entre derechos es en el campo de la inteligencia artificial.

En este orden podemos definir la inteligencia artificial como aquella disciplina que se ocupa del desarrollo de sistemas de *software* (a menudo

utilizados también en combinación con el *hardware*) que, dado un objetivo complejo, son capaces de actuar en la dimensión física o virtual, para percibir el entorno que les rodea, para adquirir e interpretar datos, razonar sobre los conocimientos adquiridos y formular decisiones, con base en las evidencias reunidas, sobre las mejores acciones a realizar para lograr el objetivo fijado, incluso en situaciones no previstas explícitamente a priori. La IA es una familia de técnicas capaces de revolucionar sectores industriales enteros, así como la interacción de los ciudadanos entre sí y con las empresas, las administraciones y la sociedad civil (**Cfm Proposte per una strategia italiana per l'intelligenza artificiale** Página web del Ministerio de Desarrollo Económico).

Como se puede apreciar la utilización de la inteligencia artificial si bien trae muchos beneficios conlleva que se puedan ver afectados diversos derechos fundamentales tales como reforzar y reproducir los prejuicios ya existentes, vulnerar los derechos de los ciudadanos a un proceso justo entre otros.

Ejemplo de estas últimas son por ejemplo los programas de predicción de actividad delictiva en los cuales puede llegar a un alto grado de discriminación racial que conlleva a acentuar aún más los perfiles discriminatorios

En efecto la capacidad de las máquinas para simular la capacidad de entender, aprender, razonar y actuar para resolver situaciones específicas o interactuar con procesos y personas, puede originar los conflictos antes mencionados y otros debido a que las máquinas son programadas por seres humanos y la selección que los mismos realizan.

En este orden juega un importante papel la ética, que en el caso de inteligencia artificial podemos definirla como un subcampo de la ética aplicada que estudia los problemas éticos que plantea el desarrollo, despliegue y utilización de la IA. Su preocupación fundamental es identificar de qué modo puede la IA mejorar o despertar inquietudes para la vida de las personas, ya sea en términos de calidad de vida o de la autonomía y la libertad humana necesarias para una sociedad democrática (**Cfm ISSN 2444-8478 • <http://doi.org/10.12795/IESTSCIENTIA.2021.i01.19> Derechos humanos, ética y transparencia algorítmica Evelyn Téllez Carvajal 372 (Grupo independiente de Expertos de Alto Nivel sobre Inteligencia Artificial, 2018, p. 13 IUS ET SCIENTIA • 2021 Vol. 7 • N° 1 • pp. 370 – 386).**

Si bien los algoritmos son esenciales para los desarrollos en materia de computación y los programas son imprescindibles en determinadas áreas como las comunicaciones y la gestión de la administración pública, relacio-

nes comerciales y muchos de los aspectos de la vida, no es menos cierto que el manejo de estos datos afecta en mayor o menor grado la privacidad. Es por eso que no sólo debe buscarse un adecuado equilibrio entre los derechos fundamentales en juego y los beneficios que pueden implicar como por ejemplo la mejora de la administración que afecta un derecho fundamental como es el derecho a una buena administración.

Ello determina que deban cada vez más regularse estos aspectos, pero mientras no se regulen corresponde a los Poderes Judiciales tratar de efectuar un adecuado equilibrio entre todos los derechos implicados evitando la imposición de discriminaciones y límites injustificados al goce de los derechos que le asisten.

Como ha expresado el Grupo de expertos en inteligencia artificial en el año 2018 todo ser humano posee un «valor intrínseco» que jamás se debe menoscabar, poner en peligro ni ser objeto de represión por parte de otros (ni de las nuevas tecnologías, como los sistemas de IA). En el contexto de la inteligencia artificial, el respeto de la dignidad humana implica que todas las personas han de ser tratadas con el debido respeto que merecen como sujetos morales, y no como simples objetos que se pueden filtrar, ordenar, puntuar, dirigir, condicionar o manipular. En consecuencia, los sistemas de IA deben desarrollarse de un modo que respete, proteja y esté al servicio de la dignidad física y mental de los seres humanos, el sentimiento de identidad personal y cultural y la satisfacción de sus necesidades esenciales. **(Grupo independiente de Expertos de Alto Nivel sobre Inteligencia Artificial, 2018, p. 15).**

Corolario de ello es que en el caso de la inteligencia artificial el derecho de cualquier individuo a conocer en detalle los criterios mecánicos que le son aplicados para decidir si posee o no un rasgo determinado, para decidir si cumple o no con un criterio predicado, es esencial. El problema se suscita entre el derecho de los afectados y la posible confidencialidad y reserva de las partes decisorias. Debemos tener presente que por un lado, existe el derecho a la información, derecho humano universal, pero también existe la posibilidad de que esta información sea reservada. Ahora bien, ello no implica que las decisiones en las cuales se basa el sistema no deban explicarse, ya que se vería altamente vulnerado el derecho a la información y el principio de transparencia, elementos claves para una adecuada defensa de los intereses en juego. No obstante cabe destacar que “El grado de necesidad de explicabilidad depende en gran medida del contexto y la gravedad de las consecuencias derivadas de un resultado erróneo o inadecuado” (**Gru-**

po independiente de Expertos de Alto Nivel sobre Inteligencia Artificial, 2018, p. 16).

Otro principio directamente implicado es el principio de equidad en el desarrollo, despliegue y utilización que conlleva a una distribución justa e igualitaria de los costes y beneficios, pero sobre todo que determinadas personas no se vean sometidas a un proceso de estigmatización o discriminación injustificadas.

Respecto de este último aspecto opera en el caso de la inteligencia artificial el principio de diversidad y de no discriminación en virtud del cual “Para hacer realidad la IA fiable, es preciso garantizar la inclusión y la diversidad a lo largo de todo el ciclo de vida de los sistemas de inteligencia artificial. Además de tener en cuenta a todos los afectados y garantizar su participación en todo el proceso, también es necesario garantizar la igualdad de acceso mediante procesos de diseño inclusivos, sin olvidar la igualdad de trato. Este requisito está estrechamente relacionado con el principio de equidad” (**Grupo independiente de Expertos de Alto Nivel sobre Inteligencia Artificial, 2018, p. 23).**

“La diversidad y la inclusión desempeñan un papel esencial al desarrollar los sistemas de IA que se utilizarán en el mundo real. Es crucial que, a medida que los sistemas de IA vayan desempeñando una mayor cantidad de tareas por sí mismos, los equipos encargados de la adquisición o del diseño, desarrollo, ensayo, mantenimiento y despliegue de estos sistemas reflejen la diversidad de los usuarios y de la sociedad en general. Esto contribuye a garantizar la objetividad y la toma en consideración de las diferentes perspectivas, necesidades y objetivos. Lo ideal es que la diversidad no sólo se materialice en los equipos en términos de género, cultura y edad, sino también de antecedentes profesionales y conjuntos de competencias”.

En este contexto, el uso de algoritmos informáticos para la toma de decisiones en el ámbito público y privado se basa en las temidas ganancias en términos de eficiencia y neutralidad. Asimismo, en muchos campos, los algoritmos prometen convertirse en la herramienta a través de la cual corregir las distorsiones e imperfecciones que típicamente caracterizan los procesos cognitivos y las elecciones realizadas por los seres humanos, destacadas sobre todo en los últimos años por una impresionante literatura económica. Pero la ganancia en efectividad y supuesta neutralidad no pueden menoscabar los derechos fundamentales respecto de los cuales se aplica.

Si bien las decisiones tomadas en algoritmo debe ser los más neutrales

y basarse en datos objetivos, no es menos cierto que su utilización implica una serie de elecciones y supuestos que no son neutrales ya que se basan en la selección de determinados valores precisos y no otros. Es por ello que la transparencia juega un rol esencial a fin de determinar la posible existencia de discriminación o que se base en datos erróneos de la realidad.

Por esta razón, la Comisión Europea conformó un grupo de expertos de alto nivel sobre la IA quienes elaboraron una serie de directrices en donde plantean que para lograr una inteligencia artificial fiable, son necesarios tres componentes: 1) debe ser conforme a la ley; 2) debe respetar los principios éticos; y 3) debe ser sólida.

Para comprobar si se cumplen estos requisitos, las directrices plantean otros requisitos esenciales: a) intervención y supervisión humana; b) solidez y seguridad técnicas; c) privacidad y gestión de datos; d) transparencia; e) diversidad, no discriminación y equidad; f) bienestar social y medioambiental; g) rendición de cuentas.

Posteriormente, en diciembre del 2018, el Consejo de Europa adopta la primera Carta Ética Europea sobre el uso de inteligencia artificial en los sistemas judiciales. Por su parte, el Parlamento Europeo emitió la resolución del 12 de febrero de 2019, sobre una política industrial global europea en materia de inteligencia artificial y robótica. así un aura de neutralidad, resultado de cálculos racionales asépticos basados en datos. No obstante, los algoritmos no tienen un carácter neutral, pues como modelos predictivos son creados sobre las bases de elecciones realizados por el ser humano que determinan los criterios utilizados en la selección de los datos, los cuales pueden tener errores o crear discriminaciones.

En primer lugar, el principio de cognoscibilidad, según el cual toda persona tiene derecho a conocer la existencia de procesos automatizados de toma de decisiones que le conciernen y, en su caso, a recibir información significativa sobre la lógica utilizando el principio de comprensibilidad, es decir, la posibilidad, de retomando la expresión del Reglamento, recibir “información significativa sobre la lógica utilizada”. En segundo lugar, el otro principio del derecho europeo relevante en la materia (pero también de trascendencia global como se utiliza por ejemplo en la conocida sentencia *Loomis vs. Wisconsin*), puede definirse como el principio de no exclusividad de la sentencia algorítmica.

La decisión automatizada “produce efectos jurídicos que no afectan o afectan significativamente a una persona”, ésta tiene derecho a que esta decisión no se base únicamente en este proceso automatizado (artículo 22

del Reg.). En este sentido, en todo caso debe existir en el proceso de toma de decisiones una contribución humana capaz de controlar, validar o denegar la decisión automática. En matemáticas e información, el modelo se define como HITL (*human in the loop*), en el que, para producir su resultado, la máquina debe interactuar con el ser humano. 15.3 En tercer lugar, del considerando núm. 71 del Reglamento 679/2016, la legislación europea se basa en otro principio fundamental, el de no utilizar el tratamiento mediante procedimientos matemáticos o estadísticos apropiados para la elaboración de perfiles, poniendo en marcha medidas técnicas y organizativas adecuadas para garantizar, en particular, que los factores sean rectificadas que supongan inexactitudes en los datos y se minimice el riesgo de error y con el fin de garantizar la seguridad de los datos personales, de forma que se tengan en cuenta los riesgos potenciales existentes para los intereses y derechos del interesado y se evite, entre otros, efectos discriminatorios contra las personas físicas por motivos de raza u origen étnico, opinión política, religión o convicciones, afiliación sindical, estado genético, estado de salud u orientación sexual, o que impliquen un uso indebido que tenga tales efectos. En este contexto, incluso frente a un algoritmo conocible y comprensible, que no es la única razón de la decisión, no debe asumir un carácter discriminatorio. En estos casos, como establece el considerando, los datos de “entrada” deben corregirse para evitar efectos discriminatorios en la toma de decisiones (**Cfm Revista de la Facultad de Derecho de México Tomo LXXI, Número 280, Mayo-Agosto 2021 DOI: <http://10.22201/fder.24488933e.2021.280-2.79666> Algoritmos y derechos humanos Haideer Miranda Bonilla**).

Como se desprende de lo analizado anteriormente la utilización de algoritmos plantea una serie de retos para el constitucionalismo actual y los derechos humanos, en particular, como el derecho a la igualdad y no discriminación, a la intimidad o *privacy*, la protección de datos personales, los derechos a la información, a la salud, al trabajo, así como al ejercicio de la participación política, entre muchos otros.

Uno de los principales retos que enfrenta esta temática es la ausencia en general de un cuadro normativo que regule las diferentes implicaciones que conlleva la inteligencia artificial. En este tema hay que reconocer los importantes esfuerzos realizados por la Unión Europea (U.E.) en donde recientemente se presentó una propuesta de regulación de la inteligencia artificial denominada “Proposal for a Regulation laying down harmonised rules on artificial intelligence” con la finalidad de abordar sus diferentes riesgos.

Esa omisión normativa conlleva a que los jueces – ordinario, constitucional y convencional – deban acudir en la resolución de los casos a los principios y valores reconocidos en la constitución y en el parámetro de convencionalidad.

En suma:

Los constantes avances tecnológicos en las diferentes áreas (informáticas, salud, etc) traen aparejado la satisfacción de nuevas necesidades y requerimientos por parte de los usuarios de los sistemas.

Y en este orden la labor del Poder Judicial es esencial a fin de determinar en que casos esas demandas encuadran dentro de la categoría de derecho fundamental y en cuales no y en su caso impulsar mediante el reconocimiento, la existencia y límites de los posibles derechos implicados.

Esta labor se ve dificultada cuando nos enfrentamos a la realidad digital continuamente cambiante.

La dificultad mayor consiste en individualizar un nuevo punto de equilibrio en la interpretación de la situación subjetiva y sus respectivos límites.

La calidad de intérprete final de la Constitución impone, a la hora de decidir la existencia o no de una compatibilidad entre la norma legal y las reglas y principios constitucionales, un criterio de prudencia, y mesura a las que debe someterse. Pero esta prudencia no puede llegar a limitar derechos fundamentales o a negar su existencia.

Si el derecho a través de la interpretación constitucional no acompaña adecuadamente las transformaciones tecnológicas y jurídicas que acontecen en una sociedad podremos llegar a cercenar derechos que gozan de protección constitucional.

No debemos perder de vista que el Poder Judicial constituye un guardián de la constitución y de los derechos fundamentales, que se involucra en la esfera de su actuación sobre el ordenamiento en el que le toca actuar, buscando siempre la posibilidad de individualizar por vía interpretativa nuevos derechos que protejan a los individuos.

Y en esa función debe tener como eje central de su accionar el individuo y en especial salvaguardar sus derechos fundamentales.

Cabe recordar que no es necesario que un derecho esté configurado expresamente en la Constitución formal para ser un derecho esencial, humano o fundamental. Y el recurrir a, la interpretación integradora del sistema de derecho permitirá desentrañar el contenido de las carencias normativas, a la que nos enfrentemos, acudiendo para ello a los valores, a los princi-

pios y a la razón histórica de la Constitución y su sistema de derecho. No debemos perder de vista que los derechos pueden no coincidir necesaria y literalmente con el elenco que surge del texto constitucional ya que pueden emerger de una fuente diversa como son los tratados internacionales.

Es importante recalcar en este orden que la expresión derechos humanos reenvía a la idea de que son derechos que pertenecen al ser humano por su condición de tal independientemente de cualquier circunstancia y reconocimiento y por consiguiente debemos analizar expresamente si el derecho sometido a nuestra jurisdicción encuadra dentro de tal hipótesis.

El constitucionalismo moderno debe tender a que exista una expansión de los derechos constitucionalmente protegidos y que los mismos converjan en una óptima convivencia e integración para una completa protección de la dignidad humana.

Es por ello que el Poder Judicial tiene una labor esencial en el impulso y desarrollo de los derechos sometidos a su jurisdicción. •

BIBLIOGRAFIA

Pino, Giorgio. *Il costituzionalismo dei diritti*, Edit Il Mulino 2017

Trattato di diritto costituzionale Tomo III, *I diritti e i doveri costituzionali*, Edit Giappichelli 2022

Pegoraro, Lucio. *Sistemi di giustizia costituzionale*, Edit Giappichelli 2019









MOMENTO MUSICAL

Clique para assistir.



Desidério Lázaro é um dos saxofonistas mais relevantes do panorama artístico nacional.

Estudou Jazz no Hot Clube de Portugal, licenciou-se em Jazz no Conservatório de Amsterdão e concluiu o mestrado em Jazz Performance na Escola Superior de Música de Lisboa (ESML).

Faz centenas de concertos por todo o país e no estrangeiro. A par da sua actividade como músico, leciona na ESML e na Universidade Lusíada de Lisboa.

Apresenta-se em variadíssimas formações que vão desde jazz mais tradicional a vertentes mais contemporâneas.

Como autor, tem apresentado os seus projetos nos principais festivais de jazz portugueses e salas de espectáculo do país. Iniciou-se nas edições discográficas em 2010 com *Rotina Impermanente*, e o seu trabalho mais recente é *Oblivion* (ASuR, 2022) tendo recebido diversas nomeações para “disco jazz do ano”.

Mário Delgado é um dos guitarristas mais solicitados nos vários quadrantes da música portuguesa, é notório o seu uso de recursos menos óbvios através de manipulação electrónica ou de outras sonoridades acústicas pouco formais, para além de se poder mover em qualquer área estilística.

Na música improvisada faz parte do trio Lokomotiv de Carlos Barretto com José Salgueiro, do Trio TGB (Tuba guitarra e bateria) com Sérgio Carolino e Alexandre Frazão e do quarteto de Carlos Martins.

Possui uma vasta discografia jazz (23 trabalhos), gravada entre 1993 e 2021, destacando-se *Taifas*, *Suite da Terra*, *Filactera*, *Labirintos*, *Evil Things*, *15 anos Lokomotiv*, *Transporte Colectivo*, *A poesia de Aldir Blanc*, *Gnosis*, *As Blue as Red*, *Entre Paredes*.

O seu disco *Filactera* foi a primeira gravação de um músico português para a editora Clean Feed.



SUPREMO
TRIBUNAL
DE JUSTIÇA

