DIURNA.

a ousadia de escrever

EDIÇÃO ESPECIAL

INTELIGÊNCIA ARTIFICIAL

FEV 2024

Nº 14

ANO 4

Número XIV

Diurna.

O Jornal Nacional dos Estudantes da Universidade Católica Portuguesa

Porto | Lisboa | Braga | Viseu

Edição | Fevereiro 2024

D.

D.

UMA EDIÇÃO. UM TEMA.
INTELIGÊNCIA ARTIFICIAL

D.

AGRADECIMENTOS

A equipa do Diurna. dedica a 14.ª Edição à Prof. Doutora Maria do Céu Patrão Neves, Presidente do CNECV, por nos ter dado a honra de uma tão rica conversa sobre os desafios éticos da Inteligência Artificial.

Ainda, aos nossos Autores, em especial, ao Prof. Doutor Henrique Sousa Antunes, que mesmo quando o tempo é escasso aceita sempre dar o seu imprescindível contributo nas nossas iniciativas sobre futuro.

Diurna.

D.

ÍNDICE

D.

Um novo regulador para a Inteligência Artificial?
Marco Caldeira e Gonçalo Mesquita Ferreira
Associado Coordenador da VdA
Associado da VdA e alumnus da Faculdade de Direito

O que podem esperar do Regulamento Inteligência Artificial os empregadores que projetem usar sistemas de IA em contexto laboral
Helena Tapp Barroso
Sócia da Morais Leitão

Desafios da Inteligência Artificial Generativa no quadro jurídico da União Europeia
Nuno Lima Luz
Associado Sénior da Cuatrecasas

Precisão e Perceção
Thomas Gaultier
Assistente Convidado da Faculdade de Direito

Paradoxos da Inteligência Artificial: Big Data Analitics e o "E" em ESG
Maria Figueiredo e Lourenço Quintão
Of Counsel da CMS e alumnus da Faculdade de Direito

O sistema financeiro como laboratório da Inteligência Artificial
Clara Martins Pereira
Assistant Professor, Durham University Law School

Inteligência Artificial e Moda: a nova era da criatividade
Maria Victória Rocha
Professora Auxiliar da Faculdade de Direito

Consulturia 4.0: a revolução do setor através de Inteligência Artificial
Pedro Penedo, Ana Rodão e José Afonso
KPMG

Entre doomers e utopians: o futuro da IA e da humanidade
Pedro Miguel Freitas
Professor Auxiliar da Faculdade de Direito

Inteligência Artificial: um dilema liberal
João Cotrim de Figueiredo
Deputado à Assembleia da República pela Iniciativa Liberal

Inteligência Artificial e política
William Hasselberger e Inês Gregório
Professor Associado e Professora Auxiliar Convidada do IEP

A Inteligência Artificial no Contexto da União Europeia
Carlos Morais Pires, António Vicente e Sofia Moreira de Sousa
Comissão Europeia (Representação em Portugal)

A aplicação da IA à Saúde: breve apontamento teórico
Maria do Céu Patrão Neves, Miguel Ricou e Inês Godinho
Presidente e Membros do Conselho Nacional de Ética para as Ciências da Vida

Inteligência Artificial e Ética
Américo José Pereira
Professor Auxiliar da Faculdade de Ciências Humanas

Inteligência Artificial e Ética
Jerónimo Trigo
Professor Auxiliar da Faculdade de Teologia

D.

D.

EDITORIAL

A edição que apresentamos, planeada num mês e realizada em dois, foi querida durante mais de um ano. Quando partilharam comigo a repentina ideia de escrever uma edição única sobre um único tema, a escolha do tema revelou-se um reflexo imediato da vontade de dar resposta à genialidade com que me deparava. Numa troca de sentenças criou-se a ideia desta concretização e, se em tudo o que fazemos pensamos sempre duas vezes, desta vez não pensámos sequer a primeira.

Isaac Asimov dizia que a ficção científica de hoje antecipa os factos científicos de amanhã. Enquanto as narrativas que lemos se tornam realidade a uma velocidade que testa a capacidade de nos julgarmos informados, os catalizadores do conhecimento intensificam a complexidade das questões e instam respostas imediatas no desígnio de evitar que se criem lacunas intoleráveis.

Numa abordagem holística, contamos com a sabedoria dos impulsionadores dos avanços tecnológicos e com a reflexão dos que permitem que esses avanços sejam possíveis - os que, como atores atrás de um palco, despendem tempo na procura de soluções justas, humanas e éticas. Numa conversa escrita, promovemos o diálogo entre os visionários audazes que ensaiam os limites do conhecimento e os que são garantes da consciência no compromisso entre o progresso e a responsabilidade.

Se Isaac Asimov tinha razão, nesta edição do Diurna. escrevemos sobre a ficção científica de ontem, com ideias audazes e pondo a tónica na inovação, numa tentativa arrojada de tentar desenhar o futuro e adivinhar os factos científicos de amanhã. Céticos à ideia de que o conseguiremos imaginar, a única certeza que temos é a de que nos teremos de reinventar.

Catarina Andrade
Diretora Nacional do Diurna.

D.

A REGULAÇÃO COMO INSTRUMENTO DE LEGITIMAÇÃO DA INTELIGÊNCIA ARTIFICIAL

Henrique Sousa Antunes

D.

No dia 1 de janeiro de 2024, a Santa Sé publicou uma Mensagem do Santo Padre Francisco para a celebração do Dia Mundial da Paz com o título "Inteligência Artificial e Paz". O texto constitui um documento fundamental para a reflexão sobre o posicionamento da humanidade perante o desenvolvimento da inteligência artificial.

Na Mensagem, a premissa das vantagens trazidas pela evolução tecnológica é acolhida. A inquietação surge, porém, no espaço de liberdade deixado à ciência. E surge com intensidade, objetando à adesão, por defeito, das utilidades associadas à inteligência artificial: «(...) A inteligência artificial deve ser entendida como uma galáxia de realidades diversas e não podemos presumir a priori que o seu desenvolvimento traga um contributo benéfico para o futuro da humanidade e para a paz entre os povos» (página 3). O enquadramento ético da atividade apresenta-se como um instrumento de validação da ciência: «O resultado positivo só será possível se nos demonstrarmos capazes de agir de maneira responsável e respeitar valores humanos fundamentais como "a inclusão, a transparência, a segurança, a equidade, a privacidade e a fiabilidade"» (página 3).

"Na Mensagem, a premissa das vantagens trazidas pela evolução tecnológica é acolhida. A inquietação surge, porém, no espaço de liberdade deixado à ciência. E surge com intensidade (...)."

Neste sentido, exige-se um controlo das práticas: «(...) Não é suficiente presumir, por parte de quem projeta algoritmos e tecnologias digitais, um empenho por agir de modo ético e responsável. É preciso reforçar ou, se necessário, instituir organismos encarregados de examinar as questões éticas emergentes e tutelar os direitos de quantos utilizam formas de inteligência artificial ou são influenciados por ela» (página 3).
A regulação desponta, assim, como uma forma de legitimação da inteligência artificial. Exige-se, mesmo, uma regulação legal, dotada da coercibilidade que adstrinja, eficazmente, o destinatário ao cumprimento das regras éticas aplicáveis. Aliás, uma regulação internacional, considerando a universalidade dos desafios colocados por esta nova era da ciência. O Santo Padre exorta «a Comunidade das Nações a trabalhar unida para adotar um tratado internacional vinculativo, que regule o desenvolvimento e o uso da inteligência artificial nas suas variadas formas» (página 8).

"A regulação desponta, assim, como uma forma de legitimação da inteligência artificial. Exige-se, mesmo, uma regulação legal, dotada da coercibilidade que adstrinja, eficazmente, o destinatário ao cumprimento das regras éticas aplicáveis."

A análise, que acompanhamos, permite esboçar uma primeira conclusão. Ao contrário do que alguns defendem, a regulação não se apresenta como um obstáculo à inovação, antes confere propósito a essa inovação. São, pois, inaceitáveis as resistências à normatividade formuladas com

assento na proteção da liberdade empresarial e na afirmação da capacidade concorrencial dos grandes blocos económicos mundiais.

"(...) a regulação não se apresenta como um obstáculo à inovação, antes confere propósito a essa inovação."

O exercício de uma função legitimadora da inteligência artificial não se basta, contudo, com o mero ato da regulação. A legitimação da inovação tecnológica depende, ainda, das finalidades prosseguidas com essa regulação.

Na Europa, assiste-se ao nascimento de um estatuto jurídico pioneiro da inteligência artificial. Com a aprovação de um quadro regulamentar na matéria, pretende a Comissão Europeia proteger o desenvolvimento, a utilização e a comercialização da inteligência artificial e, simultaneamente, salvaguardar os direitos fundamentais dos cidadãos. A perspetiva é essencialmente preventiva. Uma «inovação responsável»: «A presente proposta impõe algumas restrições à liberdade de empresa (artigo 16.º) e à liberdade das artes e das ciências (artigo 13.º), a fim de assegurar o cumprimento de razões imperativas de reconhecido interesse público, como a saúde, a segurança, a defesa dos consumidores e a proteção de outros direitos fundamentais («inovação responsável») em caso de desenvolvimento e utilização de tecnologia de IA de risco elevado. Essas restrições são proporcionadas e limitadas ao mínimo necessário para prevenir e atenuar riscos de segurança graves e possíveis violações dos direitos fundamentais» (Regulamento IA – Exposição de motivos, pág. 13).

"Na Europa, assiste-se ao nascimento de um estatuto jurídico pioneiro da inteligência artificial. (...) A perspetiva é essencialmente preventiva."

É isso suficiente? Não. A inteligência artificial é uma tecnologia disruptiva e, como qualquer movimento de descontinuidade social, constitui uma oportunidade para estabelecer novos equilíbrios na comunidade, contribuindo para a proteção e promoção dos mais desfavorecidos. A inteligência artificial pode reduzir assimetrias na prestação de cuidados de saúde, na educação, no trabalho, na justiça. De novo, revela-se a importância da Mensagem do Santo Padre: «Numa ótica mais positiva, se a inteligência artificial fosse utilizada para promover o desenvolvimento humano integral, poderia introduzir inovações importantes na agricultura, na instrução e na cultura, uma melhoria do nível de vida de inteiras nações e povos, o crescimento da fraternidade humana e da amizade social. Em última análise, a forma como a utilizamos para incluir os últimos, isto é, os irmãos e irmãs mais frágeis e necessitados, é a medida reveladora da nossa humanidade» (página 7).

D.

A regulação não deve, pois, ser, apenas, cautelar. Deve ousar intervir, dirigir, vincular a transformação tecnológica ao bem comum, e, prosseguindo este, aproveitar as virtudes da inteligência artificial como instrumento de transformação social. No Regulamento IA, ou nos espaços de complementaridade que ele preserve, e em toda a regulação geral, o legislador, europeu ou nacional, não deve temer rótulos ideológicos, cabendo-lhe fomentar uma agenda que, pondo a inteligência artificial ao serviço de todos, privilegie os esquecidos da sociedade. Em breves palavras, a inteligência artificial apresentar-se-á profundamente humana quando honrar os valores mais radicalmente humanos. E isso pressupõe a inclusão.

"A regulação não deve, pois, ser, apenas, cautelar. Deve ousar intervir, dirigir, vincular a transformação tecnológica ao bem comum, e, prosseguindo este, aproveitar as virtudes da inteligência artificial como instrumento de transformação social."

Enfim, regressa-se à Mensagem do Santo Padre: «Este processo de discernimento ético e jurídico pode revelar-se preciosa ocasião para uma reflexão compartilhada sobre o papel que a tecnologia deveria ter na nossa vida individual e comunitária e sobre a forma como a sua utilização possa contribuir para a criação dum mundo mais equitativo e humano. Por este motivo, nos debates sobre a regulamentação da inteligência artificial, dever-se-ia ter em conta as vozes de todas as partes interessadas, incluindo os pobres, os marginalizados e outros que muitas vezes permanecem ignorados nos processos de decisão globais» (página 8).

Henrique Sousa Antunes

Professor Associado da Faculdade de Direito
(Católica em Lisboa)

D.

PROIBIR A INTELIGÊNCIA ARTIFICIAL?

por Jorge Pereira da Silva

D.

Sobretudo depois da libertação do ChatGPT, pela Open AI, têm sido frequentes as manifestações de preocupação relativamente ao impacto, potencialmente catastrófico, da inteligência artificial. Nalguns casos, é mesmo aconselhada uma pausa no desenvolvimento da inteligência artificial – senão mesmo a sua proibição – até que se possam avaliar melhor as respetivas consequências.

A ideia de pausa é simplesmente absurda. Está neste momento em curso uma corrida desenfreada entre os três grandes blocos geopolíticos pelo domínio do universo digital – os Estados Unidos, a Europa e a China –, pelo que qualquer desaceleração do lado ocidental significará oferecer definitivamente a liderança ao lado chinês. A ideia de proibição é ainda mais absurda, porque até ao presente nunca foi possível bloquear o progresso.

"A ideia de proibição é ainda mais absurda, porque até ao presente nunca foi possível bloquear o progresso."

Sem qualquer paradoxo, porém, faz todo o sentido proibir o desenvolvimento e a utilização de certas aplicações específicas da inteligência artificial e, mais precisamente, daquelas que sejam incompatíveis com a dignidade da pessoa humana e com os valores humanistas que partilhamos no Ocidente. Ao contrário do que supõem os libertários tecnológicos e alguns saudosos do "maio de 68", em matéria de inteligência artificial, não é proibido proibir. Pelo contrário, a história dos direitos fundamentais está repleta de proibições.

II. Não é fácil delimitar o círculo dos direitos fundamentais afetados negativamente pela entrada da inteligência artificial na vida quotidiana das pessoas.

Uma decisão errada de um veículo autónomo pode pôr em causa a vida dos seus ocupantes e de terceiros. Um assistente pessoal, muito útil no acompanhamento de idosos dependentes, pode também ter impactos muito negativos na sua saúde. Os algoritmos que as redes sociais utilizam para identificar conteúdos ilegítimos acabam, com alguma frequência, a bloquear informações ou opiniões protegidas pela liberdade de expressão. As aplicações de análise de risco de crédito, usadas pelos bancos, produzem por vezes decisões discriminatórias dos clientes, com base em parâmetros inaceitáveis no plano valorativo. As plataformas informáticas de receção e triagem de recursos judiciais – nos países em que já são utilizadas – podem comprometer irreversivelmente o direito de acesso à justiça e, ao mesmo tempo, os direitos que o recorrente queria defender em tribunal. Enfim, as ferramentas de correção automática de provas de admissão ou de análise de candidaturas também podem comprometer o direito à educação e a liberdade de escolha de profissão.

D.

Os exemplos poderiam multiplicar-se indefinidamente, tanto quanto se têm multiplicado as aplicações da inteligência artificial. Por isso, o direito fundamental a consagrar neste domínio deve ser configurado como um direito autónomo, destinado à proteção das pessoas e da sua dignidade, e não como um conjunto de garantias de direitos que hoje integram os catálogos constitucionais. Sabemos que a proibição da pena de morte é uma garantia específica do direito à vida, que a proibição da tortura é uma garantia associada ao direito à integridade física, e que a proibição da censura é uma garantia ao serviço da liberdade de expressão. Mas as proibições de utilização de certas aplicações da inteligência artificial, bem como as fortes limitações impostas à utilização de outras – proibições e limitações prestes a ser consagradas pelo *AI Act* – não devem ser concebidas apenas como garantias dos direitos fundamentais x, y ou z. Deverão integrar o conteúdo de um direito fundamental *a se*, destinado à proteção da pessoa como um todo.

"(...) o direito fundamental a consagrar neste domínio deve ser configurado como um direito autónomo, destinado à proteção das pessoas e da sua dignidade, e não como um conjunto de garantias de direitos que hoje integram os catálogos constitucionais."

III. Na verdade, certas aplicações da inteligência artificial não tratam as pessoas como um fim em si mesmas, mas antes como um meio, um objeto, uma peça de uma engrenagem social ou de um plano que as transcende. São, por isso, contrárias à dignidade da pessoa humana, recorrendo aqui ao conhecido imperativo kantiano "trata a humanidade, que existe na tua própria pessoa, como em todas as outras pessoas, sempre simultaneamente como um fim e nunca apenas como um meio".

Em consequência, essas aplicações têm de ser juridicamente proibidas – e, porventura, a sua utilização punida. Não pode haver aqui lugar a contemplações ou matizações, à semelhança do que ocorreu com a clonagem humana, que também foi juridicamente proibida assim que a biotecnologia a tornou viável.

Já outras ferramentas da inteligência artificial, que comportam riscos elevados para os direitos das pessoas sujeitas à sua análise e às suas decisões, mas que não sejam desumanizantes, não deverão ser proibidas liminarmente. Será necessário garantir, contudo, que são escrutinadas com rigor antes de serem aplicadas pela primeira vez e que a sua utilização posterior é monitorizada continuamente.

IV. Assim, no que respeita à proibição de certas aplicações de inteligência artificial – ou direito a não ser objeto dessas aplicações –, o *AI Act* distingue entre proibições absolutas e proibições relativas.

D.

As primeiras são aplicáveis às seguintes práticas:

a) que empreguem técnicas subliminares que contornem a consciência de uma pessoa para distorcer substancialmente o seu comportamento de uma forma que seja suscetível de causar danos físicos ou psicológicos;
b) que explorem quaisquer vulnerabilidades de um grupo específico de pessoas associadas à sua idade ou deficiência física ou mental, a fim de distorcer substancialmente o comportamento de uma pessoa pertencente a esse grupo de uma forma que seja suscetível de causar danos físicos ou psicológicos;
c) que avaliem a credibilidade de pessoas singulares com base no seu comportamento social ou em características de personalidade, e em que essa classificação social conduza ao tratamento prejudicial de certas pessoas – à imagem dos denominados sistemas de créditos sociais que têm vindo a ser aplicados na China.

"(...) no que respeita à proibição de certas aplicações de inteligência artificial – ou direito a não ser objeto dessas aplicações –, o AI Act distingue entre proibições absolutas e proibições relativas."

A proibição relativa é aplicável à utilização de sistemas de identificação biométrica à distância em tempo real, em espaços acessíveis ao público, para efeitos de manutenção da ordem pública, salvo se essa utilização for necessária para alcançar os seguintes objetivos:

a) investigação seletiva de potenciais vítimas específicas de crimes, nomeadamente crianças desaparecidas;
b) prevenção de uma ameaça específica, substancial e iminente à vida ou à segurança física de pessoas singulares ou de um ataque terrorista.

V. Além da proibição de certas aplicações e do condicionamento da utilização de outras, a proteção das pessoas em face da inteligência artificial comportará ainda, muito provavelmente, duas outras dimensões: instrumentos de defesa contra a discriminação algorítmica; e a configuração de um direito a uma decisão humana. São temas que, porém, terão de ser analisados numa outra oportunidade.

Jorge Pereira da Silva

Professor Auxiliar da Faculdade de Direito
(Católica em Lisboa)

D.

NOTAS SOBRE A APROVAÇÃO DO AI ACT

(REGULAMENTO DA UNIÃO EUROPEIA SOBRE INTELIGÊNCIA ARTIFICIAL)

No passado dia 2 de fevereiro, os 27 estados membros da União Europeia deram luz verde por unanimidade, em reunião de embaixadores, à versão final do Regulamento sobre Inteligência Artificial, habitualmente designado como AI Act. Tratava-se do último momento em que as dúvidas de alguns estados sobre o seu conteúdo poderiam fazer duvidar da sua aprovação. Segue-se o trabalho técnico de fixação do texto e de tradução para todas as línguas oficiais, a aprovação em Comissão e em Plenário no Parlamento Europeu, bem como em Conselho, a nível ministerial, e a publicação no Jornal Oficial.

"Tratava-se do último momento em que as dúvidas de alguns estados sobre o seu conteúdo poderiam fazer duvidar da sua aprovação."

Trata-se provavelmente da mais importante iniciativa legislativa desta Comissão Europeia, apresentada em 21 de abril de 2021, e peça central da sua estratégia para a transformação digital.

Entre o texto inicialmente proposto pela Comissão e aquele que foi aprovado há inúmeras alterações e aperfeiçoamentos. Destaco, em especial, a introdução de regras novas para os sistemas de IA de uso geral (General Purpose AI Systems - GPAI), categoria que abrange os sistemas de IA generativa, baseados em modelos fundacionais, como os grandes modelos de linguagem (LLM), de que são exemplos o GPT4, o Bard ou o Gemini, especialmente aqueles que comportam riscos sistémicos; a introdução das Avaliações de Impacto nos Direitos Fundamentais (Fundamental Rights Impact Assessment); o alargamento da lista de usos proibidos; de certas medidas de apoio à inovação.

D.

“Entre o texto inicialmente proposto pela Comissão e aquele que foi aprovado há inúmeras alterações e aperfeiçoamentos.”

Sendo os Estados Unidos a principal potência no desenvolvimento da IA, através de grandes empresas tecnológicas (big tech), universidades e startups, estando a China em franco desenvolvimento destas tecnologias, e tendo algumas das suas empresas dimensão e valor de mercado apreciáveis, não deveria a EU centrar os seus esforços em criar condições para se tornar numa potência tecnológica comparável aos seus concorrentes, em lugar de se tornar a líder da regulação?

“(…) não deveria a EU centrar os seus esforços em criar condições para se tornar numa potência tecnológica comparável aos seus concorrentes, em lugar de se tornar a líder da regulação?”

Esta foi, recentemente, a posição assumida pela França, pela Alemanha e pela Itália. Macron foi o mais claro, aparentemente em defesa da startup francesa Mistral.

No fundo esta pergunta tem por detrás a clássica dicotomia entre regulação e inovação. Os dois conceitos são antinómicos ou compatíveis?

O AI Act classifica os usos da inteligência artificial em função do grau de risco que suscitam. Proíbe usos que considera inadmissíveis (por exemplo, a manipulação de pessoas vulneráveis, ou o reconhecimento biométrico em tempo real em espaços públicos para efeitos de aplicação da lei, embora com exceções) e impõe uma bateria pesada de obrigações para os usos qualificados como de alto risco (como os sistemas usados para avaliação de estudantes, para seleção ou recrutamento de trabalhadores, para acesso a serviços essenciais, ou para o controle das fronteiras).

“O AI Act classifica os usos da inteligência artificial em função do grau de risco que suscitam. Proíbe usos que considera inadmissíveis (…) e impõe uma bateria pesada de obrigações para os usos qualificados como de alto risco (…)”

As imposições para as empresas que usem IA neste contexto (de alto risco) são diversas, difíceis de cumprir e envolvem custos não negligenciáveis. Mas a UE não pode deixar de proteger valores e direitos fundamentais das pessoas, tendo a IA um enorme potencial para melhorar a sua qualidade de vida e, ao mesmo tempo, para criar danos consideráveis.

D.

Por outro lado, o desenho de regras claras, comuns ao espaço europeu, contribui para a criação de um mercado alargado, mais suscetível à criação de escala na estratégia de crescimento das empresas.

Acresce que o AI Act consagra regras de apoio à inovação, como a admissão de "ambientes de testagem" (regulatory sandboxes) permitindo a experimentação de usos de alto risco em condições reais controladas.

Crucial é que a regulação seja acompanhada de forte investimento na investigação e desenvolvimento, na criação de condições institucionais para as universidades e centros de investigação atraírem e reterem talento na Europa, e de incentivos significativos para as PME e startups que desenvolvem projetos neste domínio.

"Crucial é que a regulação seja acompanhada de forte investimento na investigação e desenvolvimento(...)"

Não se pode ainda esquecer que a passagem da lei à prática envolve a estrutura de governança prevista no AI Act, com autoridades nacionais necessariamente dotadas de meios adequados, para além das autoridades europeias. Em complemento dessa estrutura, e das sanções extremamente pesadas previstas, será fundamental a adoção das regras relativas à responsabilidade civil por atuação de sistemas de IA e da responsabilidade por produtos defeituosos, matérias que são objeto de projetos de diretivas ainda não aprovadas.

As empresas e outras organizações terão em princípio dois anos para se adaptarem ao AI Act. São exceção as regras quanto aos usos proibidos, que serão aplicáveis seis meses após a publicação, quanto aos sistemas de uso geral, aplicáveis após um ano, e as relativas a obrigações relativas a determinados usos de alto risco, que apenas serão aplicáveis após três anos.

Em conclusão, sem prejuízo de podermos discutir em pormenor esta ou aquela solução do AI Act, a sua aprovação é uma boa notícia para os cidadãos europeus e para a tutela dos seus direitos fundamentais. Mas só será uma benéfica para a economia e as empresas europeias se o investimento numa Europa tecnológica assente em IA for adotado como prioridade estratégia da próxima Comissão Europeia.

Luís Barreto Xavier

Professor Convidado da Faculdade de Direito (Católica de Lisboa)

D.

INTELIGÊNCIA ARTIFICIAL:

OS DESAFIOS QUE COLOCA À RESPONSABILIDADE CIVIL

por Elsa Vaz de Sequeira

D.

O desenvolvimento da inteligência artificial (IA) tem trazido, como tudo na vida, coisas boas e coisas menos boas. Se, por um lado, permite um tratamento de informação nunca antes visto, quer no tocante à quantidade de dados processados quer relativamente à velocidade e precisão com que executa semelhante operação, por outro lado, cria desafios nem sempre de fácil resolução. Um deles é justamente aquele que ora nos interessa: como lidar com os danos causados pelos sistemas dotados de IA.

"Um deles é justamente aquele que ora nos interessa: como lidar com os danos causados pelos sistemas dotados de IA".

A primeira questão que se coloca é a de saber se essas situações devem ser apreciadas à luz do princípio da imputação dos danos ou de acordo com o princípio da responsabilidade civil. Julga-se que, neste aspeto, a resposta é igual àquela que habitualmente se dá quando alguém sofre um dano. O regime-regra será o da imputação dos danos, devendo o titular do bem afetado suportar o respetivo prejuízo – ubi commoda, ibi incommoda. Este princípio cederá lugar à figura da responsabilidade civil em duas ocasiões: se o dano provier da atuação ilícita e culposa do lesante ou se o dano derivar da perigosidade inerente ao próprio sistema. O que vale por afirmar que, pelo menos em abstrato, haverá espaço para responsabilizar outrem pelo mal causado ao lesado.

"O que vale por afirmar que, pelo menos em abstrato, haverá espaço para responsabilizar outrem pelo mal causado ao lesado".

Dando um passo em frente, cabe então fazer mais duas perguntas. Concretamente, importa determinar quem pode vir a responder pelos danos causados pelos sistemas de IA e a que título

Se, num passado recente, chegou a equacionar-se a hipótese de impor esse dever ao próprio sistema de IA – reconhecendo-lhe deste modo personalidade jurídica ou algum tipo de subjetividade –, hoje em dia essa ideia parece abandonada. E julga-se que bem. Embora, de um ponto de vista estritamente técnico-jurídico, tal se afigurasse viável, de um prisma puramente pragmático, os problemas que este caminho acarretaria excedem largamente as vantagens. Nomeadamente: deveria um sistema de IA ser necessariamente dotado de um património, para fazer face a eventuais obrigações de indemnizar os danos que viesse a causar? Se sim, quem deveria contribuir para a sua formação e quem o geriria? Os ganhos obtidos estariam sujeitos a tributação? Tendo em conta que os sistemas de IA podem encontrar-se dispersos em diferentes lugares, qual seria a residência ou domicílio do mesmo? Deveria haver registo.

"Se, num passado recente, chegou a equacionar-se a hipótese de impor esse dever ao próprio sistema de IA (...), hoje em dia essa ideia parece abandonada."

D.

Não se trilhando essa via, a solução passará necessariamente por responsabilizar a pessoa que causou ilicitamente o dano ou aquela que concebeu o sistema ou retira os benefícios por ele propiciados. Na primeira hipótese, apela-se à responsabilidade por facto ilícito e culposo, enquanto nas últimas hipóteses já estaremos no âmbito da responsabilidade pelo risco. Qualquer dos trajetos se mostra espinhoso.

"Na primeira hipótese, apela-se à responsabilidade por facto ilícito e culposo, enquanto nas últimas hipóteses já estaremos no âmbito da responsabilidade pelo risco. Qualquer dos trajetos se mostra espinhoso".

Olhando ao ordenamento jurídico português, dir-se-á que os artigos 483.º e ss. do Código Civil, dedicados em especial à responsabilidade por ato ilícito, são suscetíveis de ser aplicados aos casos em análise. Isto quer dizer que existe uma previsão normativa cujo âmbito de proteção abrange estes factos. A dificuldade aqui reside tão-só na demonstração do seu preenchimento. Um sistema de IA procede do contributo de diferentes pessoas, v.g. as que concebem o algoritmo, as que introduzem os dados ou as que procedem a atualizações. Identificar o sujeito que falhou neste processo é extremamente penoso, quando não impossível. A autonomia própria do sistema também constitui um obstáculo, na medida em que torna imprevisível a sua atuação futura, inviabilizando assim quer a aferição do nexo de causalidade quer o juízo de culpa. Isto já para não falar dos casos em que se chegue à conclusão de que vários participantes no processo de criação do sistema de IA falharam, sem contudo se conseguir determinar qual das falhas foi efetivamente a causa do dano, não obstante cada uma delas ser adequada de per se para tal.

"Um sistema de IA procede do contributo de diferentes pessoas, v.g. as que concebem o algoritmo, as que introduzem os dados ou as que procedem a atualizações. Identificar o sujeito que falhou neste processo é extremamente penoso, quando não impossível".

Neste contexto, a solução mais simples parece ser o recurso à responsabilidade objetiva pelo risco. Mas, infelizmente, no quadro legal vigente não se afigura ser assim. E isto por duas razões. Desde logo, porque não vigora uma norma que o preceitue. Claro que se poderia tentar aplicar analogicamente alguma norma prevista para um caso em que a razão de decidir seja idêntica. Não raro, sugere-se a aplicação analógica dos artigos 502.º e 503.º do Código Civil, relativos aos danos causados por animais ou por veículos automóveis. A imprevisibilidade do comportamento, ali, ou a perigosidade do bem, aqui, justificariam a analogia. Crê-se, porém, que o sentido da solução não será o mais ajustado à situação concreta, pois ir-se-ia responsabilizar os proprietários, detentores ou utilizadores dos sistemas de IA, desresponsabilizando por completo aqueles que participaram na sua criação e que, devido a isso, conhecem melhor o sistema e a sua forma de funcionamento, gozando assim de uma posição de controlo sobre a respetiva perigosidade em

nada comparável com a daqueles sujeitos. Ao que acresce a circunstância de também eles, à semelhante destes, aproveitarem as vantagens proporcionados – direta ou indiretamente – pelo sistema.

"Neste contexto, a solução mais simples parece ser o recurso à responsabilidade objetiva pelo risco. Mas, infelizmente, no quadro legal vigente não se afigura ser assim."

A União Europeia tem estado atenta a estes problemas. Se, inicialmente, chegou a ponderar a implementação de um regime de responsabilidade objetiva, mais recentemente, porém, tudo indica não ser essa a via escolhida, optando antes por estabelecer como regime-regra o da responsabilidade aquiliana, acompanhado pela previsão de presunções de causalidade e/ou presunções de culpa. Paralelamente, propôs-se alterar o regime da responsabilidade do produtor para, sem sombra para dúvidas, poder contemplar as hipóteses em análise. Sendo de louvar estas iniciativas, teme-se, contudo, que não sejam suficientes.

"Se, inicialmente, chegou a ponderar a implementação de um regime de responsabilidade objetiva, mais recentemente, porém, tudo indica não ser essa a via escolhida, optando antes por estabelecer como regime-regra o da responsabilidade aquiliana (...)."

Elsa Vaz de Sequeira

Professora Auxiliar da Faculdade de Direito
(Católica de Lisboa)

D.

A (IN)APLICABILIDADE DO INSTITUTO DA RESPONSABILIDADE CIVIL AO DANO PROVOCADO POR ACTO MÉDICO PRATICADO COM RECURSO A SISTEMAS DE IA

por António de Medeiros Barbosa

D.

Nos últimos meses tenho tido a oportunidade de acompanhar várias conferências sobre o impacto da Inteligência Artificial (IA) nos vários sectores de actividade e, em especial, na área da medicina. Curiosamente, na grande maioria delas, os oradores, de forma consciente ou, quero acreditar, inconsciente, tendem a abordar esta matéria tal como se tratasse de uma realidade tão distante no tempo que só faria sentido abordá-la numa perspectiva de futuro. Esta visão, que é, aliás, transversal a diversas áreas de actividade, revela-se, porém, pouco rigorosa e algo desconforme com a realidade.

De facto, podemos afirmar com segurança que o recurso à IA no desenvolvimento da actividade médica, especialmente numa fase inicial de diagnóstico, é já uma realidade. Veja-se que no Reino Unido, um grupo de investigadores da *University College of London* juntamente com o *Moorfields Eye Hospital* desenvolveu um sistema de IA – o *RETFound* - com o propósito de identificar doenças ou sintomas de doenças que afectam a retina dos seus pacientes. Para o desenvolvimento deste sistema de IA foram submetidas cerca de 1.6 milhões de radiografias aos olhos, de forma a capacitar o sistema da informação necessária para proceder à identificação rigorosa de uma doença localizada na retina. Mas mais: este sistema de IA consegue ir mais longe sendo ainda capaz diagnosticar uma doença vários anos antes do paciente começar a ter os primeiros sintomas. É o que acontece com a Doença de *Parkinson*. Quando pensamos nesta doença, pouca relevância damos ao tecido localizado na parte de trás do olho, sobretudo, porque qualquer leigo sabe que a Doença de Parkinson afecta os movimentos e provoca tremores. No entanto, este mesmo *RETFound* consegue, com base na mera análise da retina, prever que um determinado paciente vai ter esta doença anos antes de sentir os seus primeiros sintomas.

"(...) este sistema de IA consegue ir mais longe sendo ainda capaz diagnosticar uma doença vários anos antes do paciente começar a ter os primeiros sintomas. É o que acontece com a Doença de Parkinson."

Este exemplo demonstra de forma inequívoca que, no âmbito da área da medicina, o tempo do debate do "como é que a IA vai influenciar a actividade médica?" está a atingir rapidamente o seu fim. Em abono da clarividência, começa a revelar-se algo urgente centrarmos a nossa atenção no "como é que se vai regular a intervenção da IA na prática dos actos médicos?". É que, ao reflectirmos sobre esta questão deparamo-nos com um largo conjunto de problemas a que a comunidade científica e, principalmente, a comunidade jurídica vão ter de contribuir para a sua resolução.

Aproveitando o exemplo do *RETFound*, vamos supor que um determinado paciente se dirige a uma unidade de saúde queixando-se de que tem sentido a visão ligeiramente enevoada. O médico, após recorrer ao sistema de IA, informa o paciente que não foi localizada nenhuma patologia que

D.

afectasse o seu olho. No entanto, ao contrário do diagnóstico realizado pelo *RETFound*, aquele paciente, na verdade, sofria de um grave glaucoma. Tão grave que acabou por causar cegueira umas semanas mais tarde. Tratou-se, por isso, de um erro de diagnóstico, isto é, o sistema de IA não conseguiu identificar a patologia, transmitindo ao médico informação errada sobre o estado da retina do paciente, o que acabou por provocar um dano irreversível.

Perante isto, é inevitável que surja a seguinte questão: sobre quem recai a responsabilidade do dano causado por um acto médico praticado com intervenção de um sistema de IA?

"(...) é inevitável que surja a seguinte questão: sobre quem recai a responsabilidade do dano causado por um acto médico praticado com intervenção de um sistema de IA?"

É claro que devíamos, previamente, dar resposta à questão de saber se o médico deve (ou não) ficar totalmente adstrito ao "parecer" fornecido pelo sistema de IA ou se pode, em caso de discordância fundamentada, contrariá-lo. Porém, centremo-nos apenas na questão *supra*mencionada. Ou seja: está, o ordenamento jurídico português munido de soluções para dar resposta este problema?

Aparentemente não. Repare-se, desde logo, que o princípio geral do instituto da responsabilidade civil, consagrado no artigo 483.º do Código Civil (CC), fundamenta-se na existência da culpa. Isto significa que apenas poderíamos imputar ao lesante a obrigação de indemnizar na circunstância de que este consiga compreender as consequências do seu acto ilícito e danoso na esfera jurídica do lesado e, mesmo assim, queira praticar o respectivo acto. Ora, tal circunstância não se verifica, naturalmente, na actuação dos sistemas de IA, uma vez que estes não têm vontade própria ou consciência de si e das suas acções na relação com o outro. Por esta razão, parece-nos que o instituto de responsabilidade civil, assente na ideia da culpa, tem-se revelado pouco (ou nada) adequado para dar resposta aos problemas que surgem dos danos causados pela intervenção de sistemas de IA.

Ademais, poderíamos indagar sobre a possibilidade de aplicação do regime da responsabilidade objectiva ou pelo risco, previsto no CC. Mas a questão é: será que a solução seria diferente? Será que, ao prescindir da culpa, o regime da responsabilidade objectivo revelar-se-ia juridicamente robusto resolver a problemática em análise? A resposta é, novamente negativa. Repare-se que até o regime da *"responsabilidade do comitente*, previsto 500.º do CC, que, entre nós, seria, em tese, o mais adequado, não serve na resolução do caso concreto, uma vez que, para o comitente ser responsabilizado, é sempre necessário que o acto praticado pelo comissário seja culposo. Logo, mesmo que o médico fosse o comitente e o sistema de IA o comissário, para que o primeiro fosse responsabilizado, era imperativo que o segundo praticasse um acto culposo, o que, como já vimos, não se afigura possível.

D.

"Será que, ao prescindir da culpa, o regime da responsabilidade objectivo revelar-se-ia juridicamente robusto resolver a problemática em análise? A responde é, novamente negativa."

Por tudo aquilo que acabou de ser explanado, afere-se que nem o CC não oferece uma solução juridicamente sólida para dar resposta aos problemas na área da responsabilidade civil por danos causados por sistemas de IA. Porém, se quisermos ser rigorosos, parece-nos que a solução não passa por qualquer diploma que esteja actualmente em vigor no nosso ordenamento jurídico. Desenganem-se, pois, aqueles que perante as conclusões explanas, venham agora chamar à colação a possível aplicação do Decreto-Lei n.º 383/89, de 6 de Novembro, que prevê a responsabilidade decorrente de produtos defeituosos.

"(...) o CC não oferece uma solução juridicamente sólida para dar resposta aos problemas na área da responsabilidade civil por danos causados por sistemas de IA."

Perante esta realidade, transversal a diversos Estados-membro da União Europeia, o Parlamento Europeu, tomou a iniciativa de propor um conjunto soluções jurídicas que passam, nomeadamente, pela criação de um fundo de compensação com propósito de garantir a reparação dos danos provocados pelos sistemas de IA. Com efeito, aquele paciente que foi prejudicado pelo erro de diagnóstico do *RETFound* seria indemnizado directamente por este fundo e nunca através do instituto da responsabilidade civil.

Esta solução, ainda que em fase embrionária, é, porém, reveladora de um possível caminho alternativo para fazer face às questões que suscitámos, ficando, porventura, a sensação de que o referido instituto, tal como o conhecemos, não fará parte das soluções para este problema.

Quanto a mim, afirmo: a solução não passará pelo instituto da responsabilidade civil.

António de Medeiros Barbosa

Estagiário na VdA
Alumnus da Faculdade de Direito (Católica de Lisboa)

D.

E SE A MÁQUINA DECIDIR POR NÓS?

OU OS INSTRUMENTOS DE ANTECIPAÇÃO DE RISCO DE REINCIDÊNCIA ENTRE HOJE E O FUTURO

por Pedro Garcia Marques

D.

E se a máquina decidir por nós? A questão pode bem colocar-se já hoje quando atentamos em programas computorizados de antecipação de reincidência como o COMPAS, o LSI-R, o VRAG, o UCCI e ainda o ORAS. Todos anunciados como capazes de prever o risco de reincidência de agentes condenados criminalmente pela prática de crimes. E todos lançando mão, para o efeito, de informação relativa a actuação dotada de significado estatístico, no original em inglês, actuarial-meaning statistically derived-information .

Pensados inicialmente como programas informáticos dotados de uma função de mero apoio à decisão judicial, fornecendo, numa escala de pontos, o cálculo de probabilidade de reincidência por parte de um condenado, a verdade, porém, é que, em cada um deles, se anuncia já a capacidade de, a breve trecho, a máquina estar dotada da capacidade técnica suficiente para um juízo suficientemente rigoroso sobre a certeza quanto ao risco representado pelo agente individual de vir a praticar crimes no futuro.

"(...) em cada um deles, se anuncia já a capacidade de, a breve trecho, a máquina estar dotada da capacidade técnica suficiente para um juízo suficientemente rigoroso sobre a certeza quanto ao risco representado pelo agente individual de vir a praticar crimes no futuro."

Em causa estará a possibilidade de, num futuro próximo, se contar com a possibilidade de um processo de aferição e de determinação de uma pena, na sua natureza e extensão e, com isso, na sua severidade, exclusivamente conduzido, de modo autónomo, por uma máquina.

Quando disso a máquina seja capaz a pergunta impor-se-á de modo inevitável: porque não?

Questão que se tornará tanto mais premente quanto essa máquina mostra, já hoje, uma capacidade de consideração e de análise de parâmetros e de variáveis complexos que em muito supera a capacidade humana de ajuizamento. E, sobretudo, quando, também já agora, conta ainda com uma capacidade de aprendizagem fundada na experiência, designada de deep learning, também ela muito superior àquela do correlato humano.

Quando assim fosse, caber-lhe-ia, à máquina e ao programa de que se sirva, de modo exclusivo, a determinação da pena, na sua natureza e severidade, em relação àquele agente que resulte condenado, com fundamento num juízo de mera probabilidade estatística que, de forma autónoma, determinará o risco daquele agente condenado vir a praticar crimes no futuro. Servindo aquele juízo de mera probabilidade estatisticamente calculado, como suficiente para a determinação do risco que um agente condenado representará.

D.

Probabilidade essa que se transforma, então, em certeza suficiente para justificar e, desse modo, legitimar a restrição da sua liberdade física, no modo, na forma e no tempo, em suma, nos termos em que a máquina decida e, por si mesma, determine.

"Probabilidade essa que se transforma, então, em certeza suficiente para justificar e, desse modo, legitimar a restrição da sua liberdade física (...) nos termos em que a máquina decida e, por si mesma, determine."

Mas de que juízo se trata aqui, afinal?

Sem dúvida de um juízo fundado em mera probabilidade. Fazendo para tal uso do aparato técnico de que a máquina é capaz e, bem assim, do que aprenda a cada decisão que tome, nos termos daquilo que resulte como estatisticamente calculado como possível. Remetendo, então, para as margens da improbabilidade o que, nos termos do cálculo estatístico, se apresente como tão improvável que se torna estatisticamente irrelevante.

Pouco importará, portanto, que aquele agente, individualmente considerado, não corresponda ao modelo estatístico que a máquina constrói e que, por isso, não represente efectivamente o risco que diz representar alguém como ele. O que ele seja como pessoa, na medida da diferença que, a esse nível de individual de irrepetível infungibilidade, nele se manifeste, será, no cálculo estatístico feito de regularidades probabilísticas e determinado nos termos da linearidade lógica de um raciocínio causalmente determinado, remetido para o campo da improbabilidade. Pois que, por tão raro e implausível, mais não será que estatisticamente irrelevante.

"Pouco importará, portanto, que aquele agente, individualmente considerado, não corresponda ao modelo estatístico que a máquina constrói e que, por isso, não represente efectivamente o risco que diz representar alguém como ele."

Quando assim seja, e assim será, algo se perde. E perde-se, não só para *aquele* agente condenado, mas também para a ciência que com esse agente condenado reivindica relacionar-se nos termos da obediência estrita a um quadro principiológico capaz de escorar em devida certeza e segurança e em suficiente legitimidade, cada dia, cada hora e cada minuto de uma pena que lhe seja imposta. E, com isso, algo se perde para uma ordem jurídica em que liberdade que, mais que atribuída ou concedida, a cada um, desde logo ao condenado, é reconhecida como parte inextrincável daquilo que ele *efectivamente* é. Seguramente assim será num Direito Penal próprio de um Estado de Direito democrático.

"E, com isso, algo se perde para uma ordem jurídica em que liberdade que, mais que atribuída ou concedida, a cada um, desde logo ao condenado, é reconhecida como parte inextrincável daquilo que ele efectivamente é."

D.

O que se perde, então?

Desde logo, a ideia mesma de que toda a pena concretamente determinada serve um fim preventivo especial – dirigido a cada agente condenado pela prática de um crime – e se justifica em razão desse fim, em função da estrita necessidade do que impõe em limitação de liberdade na prevenção futura da prática, por aquele agente ora condenado, de crimes futuros da mesma ou de diversa natureza.

Ora, na medida em que a tecnologia impõe o modo estritamente probabilístico do seu raciocínio, será essa pena imposta ao condenado em razão da possibilidade calculada de vir a praticar crimes no futuro, mesmo que aquele agente, *contra probabilitatem*, não represente *efectivamente* qualquer risco de reincidência.

Assim sendo, a primeira exigência a cair será a da vinculação da imposição de toda a pena, na sua natureza e severidade, à prossecução de fins preventivos. Pois que, limitação de liberdade poderá ser imposta a quem, na realidade, não representa o risco que justifica essa restrição. Nenhum risco havendo a prevenir com a imposição dessa pena.

E mais se perde.

Se em causa está um juízo de probabilidade estatisticamente calculada, será a pena imposta, na natureza e na medida da severidade, em razão, não do risco que *aquela* pessoa *efectivamente* represente, mas em razão de um risco percebido como provável, mesmo que, a consideração da *pessoa* do condenado, individualmente considerada, esse risco desminta. E o desminta pela precisa razão de que é uma pessoa e que, pelo facto de o ser, com todas as demais pessoas, partilha a novidade que a sua irrepetível *diferença* teima em transformar em imprevisível surpresa.

"Se em causa está um juízo de probabilidade estatisticamente calculada, será a pena imposta, na natureza e na medida da severidade (...) em razão de um risco percebido como provável, mesmo que, a consideração da pessoa do condenado, individualmente considerada, esse risco desminta."

E, com isso, eis que se perde a natureza pessoal do juízo que sustenta em legitimidade toda a condenação: o de culpa. Que agora será o que quer que seja, mas, *pessoal*, como deveria sempre ser, não mais será certamente.

Pelo caminho, mais fica esquecido ainda.

D.

Desde logo, o fundamento mesmo, filosófico e existencial, que traz ao nosso mundo de vida a precisa razão para que aceitemos, sem maior discussão, o que ao Direito se autoriza: a censura, a pena e a privação, dita legítima, da liberdade física. Perde-se, na verdade, o preciso indeterminismo que em cada um coloca o *modicum* de autonomia para decidir de si, projectando-se num futuro que é, em parte significa, por si construído, levando nele a marca do seu nome, o selo da sua autoria.

"(...) o fundamento mesmo, filosófico e existencial, que traz ao nosso mundo de vida a precisa razão para que aceitemos, sem maior discussão, o que ao Direito se autoriza (...)."

O futuro será produto da construção que dele cada um faça. Nada estando definido, o único destino será o da indeterminação própria de um horizonte em aberto.

Não assim, diz a máquina, pois que, para ela, cada um será, apenas e só, o que o cálculo linear determinar como provável. O risco que essa probabilidade apontar será a definição mesma da pessoa e do que nela importa neutralizar.

Se és apenas o que fizeres, o que fores será aquilo que a estatística determinar, como se de certeza se tratasse.

Se arrependimento ou epifania, sorte ou acaso, introspecção ou metamorfose fizerem a sua entrada numa vida que, enquanto humana, é sempre imprevisível e surpreendente, nada disso importa à máquina que sujeita a liberdade de cada um, desde logo, do condenado, ao determinismo mecanicamente calculado.

No entanto, se do indeterminismo desistimos, desistiremos também do peso que a dúvida deve exercer na ponderação da prova produzida em tribunal e no modo como ela se deve reflectir no arguido, desde logo, na sua liberdade e na necessidade da sua preservação.

Pois que, satisfeita a máquina com a probabilidade, torna-se ela co-essencial à dúvida.

Aceite que, onde essa dúvida exista, nem por isso se impede a decisão no sentido de existência do risco de reincidência, será essa decisão tomada mesmo que, em face da dúvida, risco haja de erro. De erro sobre se, de facto, o condenado representa o risco de reincidência que se calcula quanto a ele como existente, porque provável. O erro, em suma, de se lhe impor uma pena para prevenir um risco que ele *efectivamente* não representa.

D.

Quando assim seja, dúvida e risco de erro sobre aquele juízo remetidos se tornarão para o lugar de *custo* do sistema. Um *custo* que em nada impede a máquina de funcionar e de impor *a sua lei,* mas que será pago por alguém na forma de liberdade indevidamente restringida. Pelo meio custando também o lugar, o papel e a função que a *dúvida* deveria desempenhar no Direito: aquela de *in dubio por reo.*

"Pelo meio custando também o lugar, o papel e a função que a dúvida deveria desempenhar no Direito: aquela de in dubio por reo."

De um *dubio* que, afinal, neste admirável mundo novo, o será, as mais das vezes, *contra reo.*

Neste mundo, *admirável* e novo, importarão, sim, os factos que, objectivos, contribuirão para a extrapolação estatística de comportamentos, nos termos de uma regularidade linear e probabilística. E para a máquina, dir-lhe-á mais o lugar onde o condenado habita, as habilitações académicas que tem, o nível de rendimento que aufere, o estar ou não empregado, a sua história de adição e, sim, a sua etnia, a sua história familiar, o meio social e a classe a que pertence.

Tudo contribuindo, no entanto, para um risco de *bias* que, à identificada indiferença do que de humano existe em cada um, acrescenta, em desconsideração e em desrespeito, o peso do preconceito e da assimetria de reconhecimento. Contribuindo a indiferença em relação a esse bias para perpetuar um e outro, nunca para os corrigir. E, com isso, do mesmo passo contribuindo para o perpetuar do que um e outro representam: a negação da dignidade social ao todos devida e com que, amiúde, se viu o agente, desde o berço, sobrecarregado.

Não considerando a favor do condenado a *dúvida*, nunca rejeitando a mera probabilidade, antes a tornando acriticamente aceite, perpetuando o preconceito que, mais do que à pena, condena o agente *às margens* de uma comunidade que nas margens tantas vezes o manteve, a privação de liberdade imposta pela máquina e pelo programa de que fará uso tornar-se-á na manifestação visível do que sempre faltará no cálculo cibernético determinado por ela promovido: a de que *aquele* sobre *quem* a máquina decide é *alguém* sobre o qual ela, não apenas nada sabe, como alguém de quem não quer saber.

Quem está ali afinal? Não interessa.

No final, o que se perde é própria *pessoa.*

Pedro Garcia Marques

Professor Auxiliar da Faculdade de Direito (Católica de Lisboa)

D.

INTELIGÊNCIA ARTIFICIAL ("IA") E CONTRATAÇÃO

A IA é um fenómeno transversal a várias disciplinas científicas. No Direito dos Contratos, a IA manifesta-se nas fases da negociação e execução. São hoje conhecidos "sistemas de IA" que analisam clausulados contratuais, numa lógica de custo-benefício, socorrendo-se de parâmetros quantitativos (range of prices & volumes), qualitativos (range of quality) e de outras coordenadas factuais. Entre as vantagens associadas é usual referir-se a celeridade, a simplificação, a eficiência e a racionalidade dos processos negociais. Noutro campo, o "Smart Fridge" promete agilizar as exigências do "day-to-day", ao colocar encomendas de produtos com autonomia, ponderando os padrões de consumo verificados.

"São hoje conhecidos "sistemas de IA" que analisam clausulados contratuais, numa lógica de custo-benefício, socorrendo-se de parâmetros quantitativos (...), qualitativos (...) e de outras coordenadas factuais"

O recurso à IA na contratação evidencia-se, em particular, nos seguintes cenários:

- Contrato celebrado entre um contraente com a natureza de pessoa jurídica e um sistema de IA;
- Contrato celebrado entre dois sistemas de IA (contrato "M2M" = "machine-to-machine" ou "self-driving contract")

D.

A contratação com recurso a sistemas de IA – desassistidos, como tal, de "querer negocial" – põe em causa a suficiência e a adequação de figuras e regimes vigentes assentes no elemento da vontade:

- Em primeiro lugar, o conceito de contrato como "acordo de vontades" (cfr. artigo 232.º do Código Civil – "C.C.");
- Em segundo lugar, as noções de "parte" e de "representante voluntário", em razão das dúvidas quanto ao centro de imputação dos efeitos jurídicos;
- Em terceiro lugar, os regimes jurídicos sobre patologias negociais, como os vícios na formação da vontade e as divergências entre a vontade e a declaração (cfr. artigos 240.º a 257.º do C.C.);
- Em quarto lugar, o regime da responsabilidade civil subjectiva por factos ilícitos e da responsabilidade civil pré-contratual: não podendo ser sindicada a conduta do sistema de IA com base num juízo de censura, pode ficar prejudicada a indemnização dos danos causados, v.g., pela prestação de informações falsas ou parciais (cfr. artigo 227.º do C.C.);
- Em quinto lugar, o regime a observar na interpretação do contrato, considerando a insusceptibilidade de ponderar, como critério primário, o conhecimento sobre a "vontade real" dos contraentes (cfr. artigo 236.º, n.º 2 do C.C.) e as dificuldades em determinar o sentido de conceitos ambíguos (cfr. artigo 237.º do C.C.).

Como corolário do que se referiu, a contratação "M2M" permite afirmar uma tendencial "crise" de directrizes gerais de Direito dos Contratos, como: (i) a autonomia privada e a liberdade contratual (cfr. artigo 405.º do C.C. – com a supremacia do determinismo e o consequente esvaziamento do poder de, no exercício de um acto de vontade livre e esclarecida, decidir como se contrata, com quem se contrata e se se contrata) e (ii) a estabilidade dos contratos (cfr. artigo 406.º, n.º 1 do C.C. – que releva, entre as causas de modificação ou de cessação de vigência do contrato, o "mútuo consentimento").

"(...) a contratação "M2M" permite afirmar uma tendencial "crise" de directrizes gerais de Direito dos Contratos".

Esta "nova forma de contratação" comporta riscos relevantes, a saber:

- A standardização de comportamentos negociais e de clausulados, em resultado dos exercícios de comparações sucessivas e correlações lógicas;
- O desequilíbrio negocial, nas hipóteses em que um dos contraentes é um sistema de IA, com a natureza de "super parte";
- A incerteza quanto ao momento da vinculação negocial, em concreto, quanto à existência de uma proposta válida e eficaz;

D.

- A opacidade negocial (em virtude do denominado "black box effect") relativamente aos elementos ponderados na decisão de contratar;
- A assimetria informativa, que se evidencia no confronto entre as características funcionais do sistema de IA (considerando as competências qualificadas no plano da recolha, tratamento e análise de dados) e as características e competências médias de um sujeito jurídico;
- O recurso a conceitos "neutrais" e sem o rigor técnico-jurídico adequado, com as consequentes dúvidas interpretativas;
- A debilidade de meios de tutela jurídica, em hipóteses de contratação sem esclarecimento ou com erro;
- A tendencial rigidez dos processos de formação e de execução contratual;
- A incerteza quanto ao âmbito da tutela patrimonial dos credores, em especial, no que respeita ao figurino do "devedor" da prestação e à suficiência da garantia patrimonial do crédito.

Diagnosticado o problema e identificados os principais riscos da contratação com sistemas de IA, cabe, agora, ao Direito definir a "terapêutica" adequada, isto é, aprovar um núcleo fundamental de regras harmonizadas e de "guiding principles" que integrem o "futuro Direito dos contratos inteligentes".

"Diagnosticado o problema e identificados os principais riscos da contratação com sistemas de IA, cabe, agora, ao Direito definir a "terapêutica" adequada (...)."

Ana Filipa Morais Antunes

Professora Auxiliar da Faculdade de Direito (Católica em Lisboa)

D.

INTELIGÊNCIA ARTIFICIAL E DIREITO DE AUTOR

por Tiago Bessa

D.

Gosto de pensar que quando Arthur C. Clarke terminou a escrita de «2001: odisseia no espaço», em 1968, estaria longe de imaginar que o mítico supercomputador que batizou como HAL 9000 (acrónimo de *Heuristically programmed ALgorithmic computer*), o cérebro e sistema nervoso da nave *Discovery,* que visava determinar o lugar da humanidade no universo, poderia ser uma realidade a breve trecho. No entanto, volvidos pouco mais de 50 anos, o que está em páginas da mais bela ficção científica já escrita, está prestes a concretizar-se, não como ameaça à espécie humana, mas sim como a emergência de uma Inteligência Artificial (IA) capaz de rivalizar com a mente humana e surgimento de uma ferramenta essencial para a evolução da nossa sociedade.

A IA está ao virar da esquina e os seus impactos ainda não estão totalmente desvendados. Há, no entanto, alguns campos onde as consequências dessa confluência já são visíveis. É o caso da criatividade intelectual, considerado por muitos como um atributo e reduto inescapável da humanidade, e da forma de proteção dessa expressão criativa, o chamado direito de autor.

"A IA está ao virar da esquina e os seus impactos ainda não estão totalmente desvendados."

A interseção do direito de autor com a IA levanta problemas delicados, que, para efeitos de discussão, podem ser divididos em duas temáticas distintas. De um lado, os problemas relativos ao input que é necessário para efeitos de treino e desenvolvimento de capacidade de aprendizagem dos modelos de IA. Do outro, os problemas relativos à eventual proteção que deve ser concedida ao output gerado por estes modelos.

Quanto ao primeiro conjunto de problemas, vale a pena recordar que no início de 2024, Sam Altman, CEO da OpenAI, a empresa por trás do ChatGTP, referiu o seguinte: "*it would be impossible to train today's leading AI models without using copyrighted materials*."

Esta afirmação evidencia que o desenvolvimento de modelos de IA, em especial os modelos de IA generativa, dependem substancialmente do acesso e utilização de conteúdos protegidos por direito de autor, como músicas, vídeos ou textos. Sem acesso e utilização destes elementos não é possível *educar* esses modelos e assim permitir o desenvolvimento de novos produtos e serviços.

A questão que aqui surge é saber se o referido acesso e utilização está dependente da autorização dos titulares de direitos. O direito de autor é um tradicional direito de exclusivo, que tipicamente atribui um monopólio na sua exploração económica, o que significa que, como regra, o uso de obras depende de uma autorização prévia dos respetivos titulares de direitos. Só assim não é quando estão previstos determinados limites ou exceções.

D.

Antevendo-se a dificuldade em obter autorizações prévias dos diferentes titulares de direitos, tem-se procurado enquadrar tal uso num limite ou exceção ao direito de autor. A este respeito, tem surgido o tema da prospeção de textos e dados, que é basicamente uma forma de permitir a mineração ou extração de informações disponíveis na Internet para assim treinar modelos de IA. O problema é que essa utilização para efeitos comerciais (para efeitos de investigação científica é diferente) tem alguns limites e pode estar dependente de saber se os titulares de direitos autorizaram ou não essa utilização, o que também implica saber qual o catálogo de obras utilizada no treino de modelos de IA.

O futuro regulamento de IA procura dar resposta a este problema referindo que as entidades que exploram modelos de IA generativa devem respeitar o direito de autor e devem publicar informação detalhada sobre a informação utilizada para treinar os seus modelos, o que pode ser depois utilizado pelos titulares de direitos para perceber se as suas obras estão ou não a ser licitamente utilizadas.

Enquanto se espera por saber qual o sucesso deste novo sistema, que, em boa verdade, deixa muitas dúvidas sobre a sua aplicação prática, vão-se sucedendo os casos em Tribunal, como o recente do New York Times, por violação de direito de autor. É seguro que a procissão ainda está no adro.

O segundo conjunto de problemas diz respeito ao output gerado pelos modelos de IA. Os conteúdos gerados por IA estão a fazer avançar o acervo cultural e educacional da nossa sociedade, permitindo uma nova forma de expressão da liberdade criativa dos artistas, mas também o surgimento de caminhos criativos alternativos e autónomos. Existem numerosos exemplos da capacidade da IA para criar obras, quase que emanando um espírito criativo próprio que não depende necessariamente das instruções dadas pelo agente humano.

“Os conteúdos gerados por IA estão a fazer avançar o acervo cultural e educacional da nossa sociedade, permitindo uma nova forma de expressão da liberdade criativa dos artistas, mas também o surgimento de caminhos criativos alternativos e autónomos."

A questão que aqui se levanta é saber se tais obras devem ou não ser protegidas e, em caso positivo, quem deve ser o respetivo titular. Esta questão vai ao âmago do direito de autor que surgiu, em termos sócio-filosóficos, pelo menos na Europa, como uma forma de tutelar a personalidade do criador intelectual - a expressão criativa é protegida porque ela reflete a personalidade do seu criador. Ora, isso não existe na *criação* por IA, que também não precisa de monopólios ou outros incentivos económicos para continuar a criar, o que tipicamente é exigido como forma de incentivar a criação intelectual.

D.

Claro está que obras em que a IA apenas seja utilizada como ferramenta para exponenciar a criação não devem gerar questões de proteção. Nestes casos, a IA é uma ferramenta, como qualquer outra, que é utilizada pelo seu criador intelectual. Mais questões surgem quando é o modelo de IA que decide as opções ou faz as escolhas criativas, sem que se possa imputá-las a um criador humano. A este propósito, existem diversas centenas de exemplos que poderiam ser citados e a lista não para de engrossar.

Na generalidade dos países, o quadro legal apenas admite proteção quando existe uma obra que possa ser imputada a um criador humano, pelo que, em regra, as obras criadas por modelos de IA pertencem ao domínio público e podem ser utilizadas sem autorização. Poder-se-á questionar se a situação se deve manter assim ou se o sistema jurídico deve avançar e enquadrar as novas formas de realidades criativa (até já se aludiu à possibilidade de expandir o conceito de personalidade jurídica), até porque embora os modelos de IA não sejam sensíveis aos típicos incentivos gerados pela proteção intelectual, as entidades que os desenvolvem seguramente que são.

"(...) em regra, as obras criadas por modelos de IA pertencem ao domínio público e podem ser utilizadas sem autorização."

O Direito é lento, mas tem ou encontra sempre soluções muito ricas que permitem e tem permitido enquadrar os principais desenvolvimentos tecnológicos das últimas décadas. Existem, por isso, diferentes soluções que tem vindo a ser apontadas, como um direito conexo ou direito *sui generis* das entidades que desenvolvem os referidos modelos e que permita tutelar o seu investimento. Mas o tema não está, por ora, estabilizado e permanece em acesa discussão.

"O Direito é lento, mas tem ou encontra sempre soluções muito ricas que permitem e tem permitido enquadrar os principais desenvolvimentos tecnológicos das últimas décadas."

Ao contrário de David Bowman, herói central de «2001: odisseia no espaço», que não encontra as respostas planeadas na expedição *Discovery*, a interseção entre IA e direito de autor precisa cada vez mais de respostas urgentes. Disso dependerá o salto quântico em evolução tecnológica prometido pela IA e a sua adoção por todos nós.

Tiago Bessa

Sócio de Comunicações, Proteção de Dados e Tecnologia da VdA

D.

PATENTES E INTELIGÊNCIA ARTIFICIAL:

NADA DE NOVO DEBAIXO DO SOL?

O direito de patentes tem um objetivo fundamental: alinhar incentivos essencialmente (embora não exclusivamente) económicos, de modo a estimular a criação e divulgação de soluções técnicas para problemas técnicos. Em contrapartida da partilha de inovações tecnológicas com o resto da sociedade, ao titular da patente é atribuído um exclusivo sobre a invenção, que pode durar até um máximo de 20 anos a contar da data do pedido. É precisamente através da concessão deste exclusivo – que mais não é do que o direito de, durante um período limitado de tempo, explorar a invenção em condições monopolísticas – que o direito de patentes procura incentivar a inovação tecnológica.

"É precisamente através da concessão deste exclusivo – que mais não é do que o direito de, durante um período limitado de tempo, explorar a invenção em condições monopolísticas – que o direito de patentes procura incentivar a inovação tecnológica."

Sem menosprezar a capacidade criativa de que possam ser dotados outros animais, árvores, fungos, bem como a própria natureza enquanto sistema – capazes que são de gerar soluções verdadeiramente engenhosas para os desafios que enfrentam –, o inventor é, por definição, um ser humano.

Os sistemas computacionais têm vindo a ganhar capacidades exponencialmente melhores, permitindo avanços científicos assinaláveis, nomeadamente no setor da investigação médica e biotecnológica. Nos casos em que esses sistemas são usados como meras ferramentas, não haverá dúvidas de que, para o direito de patentes, os inventores serão os seres humanos que exercem controlo sobre o sistema. Diríamos que estes são, em terminologia Hartiana, easy cases.

D.

Não obstante, nos últimos anos têm vindo a ser utilizados sistemas capazes de gerar invenções de forma autónoma, sem que o processo inventivo seja verdadeiramente controlado por um ser humano. O "pensamento" é "da máquina". Claro que continua necessariamente a haver intervenção humana a montante – na conceção e desenvolvimento do sistema – e a jusante – na análise e validação dos resultados. Parece-nos que essa intervenção humana pode ser suficiente para que, do ponto de vista do direito de patentes, continuemos a qualificar estes seres humanos, que controlam o sistema, como inventores.

"Não obstante, nos últimos anos têm vindo a ser utilizados sistemas capazes de gerar invenções de forma autónoma, sem que o processo inventivo seja verdadeiramente controlado por um ser humano."

Recentemente, alguns autores têm vindo a argumentar em sentido diferente, defendendo que, uma vez que a verdadeira atividade inventiva decorre "no interior" do sistema de IA, é este – e não o ser humano que o opera ou que o concebeu – que deve ser reconhecido como inventor. Foram inclusive proferidas várias decisões judiciais sobre o tema, com destaque para as decisões relativas ao sistema DABUS desenvolvido por Stephen Thaler (decisões essas, em geral, desfavoráveis à qualificação do sistema de IA como inventor). Como é sabido pela generalidade dos leitores do Diurna., não existem direitos sem sujeito, sendo o primeiro obstáculo ao reconhecimento desta qualidade a ausência de personalidade jurídica desse suposto "inventor-máquina". Por outro lado, os objetivos fundamentais do sistema de patentes, como incentivo ao progresso técnico e económico, continuam a ser dirigidos aos seres humanos.

Apesar da tendência para a antropomorfização dos sistemas de IA, não se deve perder de vista que o que está em causa são (ainda) ferramentas apenas, desenvolvidas e operadas por seres humanos, idealmente para benefício da humanidade. Enquanto assim for, cremos que os quadros tradicionais do direito de patentes não correm o risco de ser abalados pela Revolução da IA. Recorrendo à conhecida citação do Eclesiastes – utilizada num caso clássico de patentes norte-americano – diríamos que, por enquanto, não há nada de novo debaixo do sol...

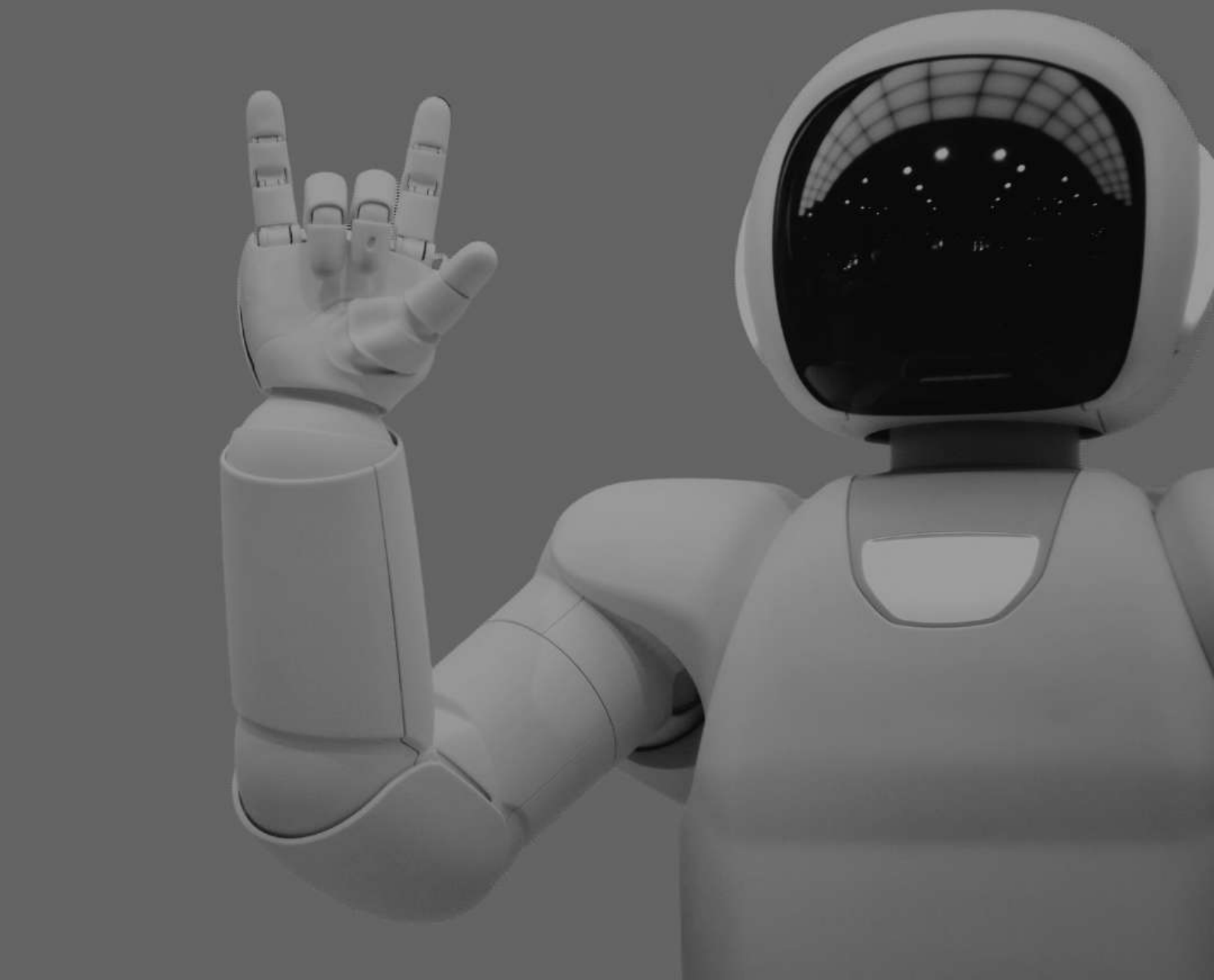
Tito Rendas e Nuno Silva

Professores Auxiliares da Faculdade de Direito (Católica de Lisboa e Católica do Porto)

D.

IA E A AMEAÇA AO LIVRE-ARBÍTRIO

Filipa Calvão



D.

A inteligência artificial (doravante, IA) apresenta-se hoje como uma inevitabilidade, não obstante o sentido e o ritmo da sua evolução futura serem ainda uma incógnita. Certo é, porém, que, ao lado das vantagens que as tecnologias de IA, máxime os modelos de autoaprendizagem, trazem em diferentes domínios e, sobretudo, na gestão do risco e da incerteza, também colocam a Humanidade perante um desafio imediato. Não me refiro ao (por ora distante) risco de domínio da IA sobre a inteligência humana, mas ao risco de eliminação ou redução substancial da autonomia individual e do livre-arbítrio, com direta repercussão nas sociedades democráticas.

Consabidamente, as tecnologias de autoaprendizagem dependem da análise maciça de dados – entre os quais se encontram dados que, quando não identifiquem diretamente as pessoas a que dizem respeito, permitem a sua identificação (dados pessoais) –, para, por via da deteção de padrões e de construção de perfis, inferir as probabilidades de ocorrência de comportamentos humanos ou de eventos diversos. Ora, o alargamento crescente das fontes da informação, abrangendo praticamente a totalidade dos seres humanos, e a granularidade da informação pessoal possibilitam um retrato detalhado de cada indivíduo e a rastreabilidade das suas ações e relações, o que condiciona o gozo das liberdades fundamentais e afeta o desenvolvimento da personalidade e da identidade de cada um (modelada em função do comportamento da maioria ou de recomendação comportamental proveniente de um sistema algorítmico).

"(...) o alargamento crescente das fontes da informação, abrangendo praticamente a totalidade dos seres humanos, e a granularidade da informação pessoal possibilitam um retrato detalhado de cada indivíduo e a rastreabilidade das suas ações e relações, o que condiciona o gozo das liberdades fundamentais (...)".

Constituindo, por isso, poderoso instrumento de vigilância e de modelação do pensamento e do comportamento humanos, concentrado nas mãos de um número reduzido de grandes grupos económicos ou de alguns Estados, a IA representa uma ameaça imediata para o que nos identifica como seres livres e pensantes: o nosso livre-arbítrio; e, assim, para o fundamento das sociedades democráticas – o conjunto das vontades individuais dos membros da sociedade, livremente formadas e manifestadas. Acresce que, por força da opacidade do processo de produção de resultados ínsita a alguns modelos de IA (redes neuronais profundas), se torna impossível ou, pelo menos, muito difícil, concretizar uma avaliação prévia do seu impacto nas liberdades e dimensões pessoais fundamentais (v.g., privacidade). Com isto fica prejudicada qualquer ponderação adequada entre esse risco e os interesses cuja satisfação justifica, prima facie, a utilização das novas tecnologias (não só quando tais interesses são reconhecidos por lei como interesses gerais ou públicos, mas também quando correspondem a interesses legítimos de privados apresentados como instrumento de realização do progresso e do bem comum). Todavia, abdicar dessa

D.

avaliação e ponderação em concreto acaba por traduzir a renúncia à matriz humanista que caracteriza as sociedades europeias.

"(...) fica prejudicada qualquer ponderação adequada entre esse risco e os interesses cuja satisfação justifica, prima facie, a utilização das novas tecnologias".

Face a estes desafios, sucedem-se as questões quanto à concretização da função do Direito neste novo contexto tecnológico. Como pode o Direito cumprir a sua função de conformação da vida em sociedade e de orientação das condutas humanas, quando o objeto da sua regulação está em constante e acelerada evolução? Como se pode garantir a regulação jurídica da utilização de uma realidade que o ser humano desconhece, dada a opacidade de alguns modelos de IA? De que modo se assegura o respeito pelo princípio da dignidade humana ou por outros princípios estruturantes, como o da proporcionalidade, na restrição de direitos fundamentais determinada por tais tecnologias? Que mudanças na estrutura e linguagem normativas têm de ocorrer para se regular a utilização destes modelos de IA e garantir o seu controlo (humano ou tecnológico) eficaz? Estas e outras questões reclamam a atenção redobrada de todos, e em particular dos juristas, de modo a, no contexto de um diálogo interdisciplinar e supranacional, encontrarmos um caminho regulatório capaz de assegurar o respeito pelos valores humanistas que identificam a nossa sociedade.

Como pode o Direito cumprir a sua função de conformação da vida em sociedade e de orientação das condutas humanas, quando o objeto da sua regulação está em constante e acelerada evolução?

Filipa Calvão

Presidente da Comissão Nacional de Proteção de Dados
Professora Associada da Faculdade de Direito

D.

A IA NO MUNDO DO DIREITO, POR "MARES NUNCA ANTES NAVEGADOS"

REGULAÇÃO DA IA: RISK BASED APPROACH

D.

Em dezembro de 2023, o Parlamento Europeu e o Conselho chegaram a consenso sobre as traves-mestras do regime europeu em matéria de inteligência artificial (IA), com o objetivo de identificar e gerir os riscos associados aos sistemas de IA (risk based approach). Este modelo de regulação assente em níveis de risco (risco inaceitável, elevado e limitado) terá um vasto impacto no mundo do Direito.

No plano dos riscos inaceitáveis, alguns sistemas de IA considerados uma ameaça para as pessoas e proibidos à luz dessa regulamentação admitirão algumas exceções para fins de aplicação da Lei. Em particular, certos sistemas de identificação biométrica serão permitidos em relação a alguns crimes graves asseguradas determinadas garantias.

"No plano dos riscos inaceitáveis, alguns sistemas de IA considerados uma ameaça para as pessoas e proibidos à luz dessa regulamentação admitirão algumas exceções para fins de aplicação da Lei".

Ao nível dos sistemas de IA que representam um risco elevado por afetarem negativamente a segurança ou os direitos fundamentais, a aplicação da Lei e a assistência na interpretação e aplicação da Lei serão áreas específicas objeto de tutela, com imposição de exigências em relação, por ex., à qualidade dos dados, ao registo de atividades e informação aos utilizadores, à supervisão humana e cibersegurança.

Os demais sistemas de IA de risco limitado poderão estar sujeitos a requisitos menos exigentes de transparência perante os utilizadores e eventuais regras de conduta de adesão voluntária.

Muitas dúvidas persistem sobre como a regulação e aplicação das regras e princípios gerais dos vários ramos do Direito e de diversas áreas de especialização lidarão com estes riscos (por ex., nos domínios dos direitos fundamentais, responsabilidade penal e civil, propriedade intelectual e proteção de dados).

Aplicação da IA: "trabalho de equipa" entre a tecnologia e o ser humano

A experiência prática dos últimos anos tem demonstrado que o recurso à IA constituirá uma oportunidade e um desafio significativo no mundo jurídico, cujo sucesso dependerá do equilíbrio entre, por um lado, a preservação dos valores éticos e direitos fundamentais que a regulamentação referida pretende preservar e, por outro, a capacidade de inovar, gerir a mudança e desenvolver a transição digital dos vários atores do mundo jurídico (como sejam legislador, tribunais, autoridades policiais, supervisores, advogados).

D.

O uso da tecnologia pelos legisladores, tribunais, autoridades policiais e supervisores poderá trazer ganhos substanciais de eficiência, mas constitui uma das dimensões de maior desafio de modo a salvaguardar a confiança na Justiça e a proteção dos direitos fundamentais e dos valores éticos.

"O uso da tecnologia pelos legisladores, tribunais, autoridades policiais e supervisores poderá trazer ganhos substanciais de eficiência (...)".

Por sua vez, na advocacia tem-se vindo (e continuar-se-á cada vez mais) a recorrer a novos métodos de trabalho, da perspetiva quer da eficiência dos processos internos quer das soluções para os clientes (também eles perante processos de transformação digital nos mais diversos setores de atividade).

Estudos recentes estimam que no futuro cerca de 40% do trabalho jurídico pode ser automatizado com recurso a IA.

A título de exemplo, o recurso à IA pode trazer e tem vindo a trazer nos últimos anos contributos em áreas como:

- Na denominada advocacia de negócios, o recurso a ferramentas de IA tem evidenciado ganhos de eficiência na análise jurídica e legal due diligence de volumes documentais significativos (desde documentos mais standardizados, como licenças ou documentação imobiliária, passando por certos contratos comerciais) e na produção automatizada de documentos (por ex. de contratos e registos), em ambos os casos com uma utilização supervisionada;

- Em matéria de investigação, contencioso e gestão de risco, destacam-se as ferramentas de electronic discovery que permitem a análise preventiva de dados em massa, em períodos temporais curtos (por ex., nos domínios jus-concorrenciais, penais, contraordenacionais ou regulatórios).

De modo crescente, a IA generativa pode exponenciar o contributo destas ferramentas no trabalho jurídico, sobretudo para lidar com tarefas repetitivas, pesquisas jurídicas e grandes quantidades de dados e para automatizar processos, exponenciando igualmente os riscos associados ao recurso à IA.

Face às oportunidades e riscos envolvidos, os contributos da IA com impacto positivo no mundo do Direito dependem do "trabalho de equipa" entre as capacidades da tecnologia e o talento e a ética do ser humano, por duas ordens de razões. Desde logo, o papel dos vários atores do mundo jurídico

D.

é essencial para assegurar a qualidade, a segurança e o cumprimento rigoroso de deveres deontológicos no mundo do Direito, orientados por um quadro regulatório que se espera promova o equilíbrio entre a inovação e a ética. Adicionalmente, a aplicação da IA resulta do ensino humano, através de processos de machine learning, pelo que o papel dos vários atores do mundo jurídico passará também a ser o desenvolvimento, o ensino e o aperfeiçoamento destes modelos.

"(...) os contributos da IA com impacto positivo no mundo do Direito dependem do "trabalho de equipa" entre as capacidades da tecnologia e o talento e a ética do ser humano, por duas ordens de razões(...)".

A academia não passará ao lado destes desafios e terá um papel decisivo neste caminho, em três principais frentes: o ensino e a produção científica nas faculdades de matérias relacionados com a regulação da IA; a crescente multidisciplinariedade e cooperação entre a Tecnologia e o Direito; e o suporte a um movimento de transformação cultural na prática jurídica inerente ao recurso a ferramentas de IA (incluindo no contexto de sandboxes para testar e implementar ferramentas de IA no setor jurídico).

A adoção crescente da IA de modo transversal na vida dos cidadãos, pelo Estado e pelo setor jurídico trará, assim, "novos mundos ao mundo" por "mares nunca antes navegados".

Magda Viçoso
Sócia da Morais Leitão

D.

UMA MAÇÃ NO ESCURO | MIGUEL MOTA

D.

"Ela, a máquina, está começando também a criar problemas que desorientam e engolem o homem. A máquina continua crescendo. Está enorme. A ponto de que talvez o homem deixe de ser uma organização humana. E como perfeição de ser criado, só existirá a máquina.", escrevia, em 1970, Clarice Lispector, muito antes de poder testemunhar veículos que conduzem, imagens que são delírios da máquina, ou deepfakes do Biden a jogar Counterstrike com o Trump. Mesmo sem conseguir adivinhar explicitamente uma inteligência artificial, a ilustre carioca esboçou, na sua crónica, um sentimento de 2024: um espanto mudo com o infinito potencial duma inteligência que ultrapassa a nossa, seguido da realização de que o Shangri-La que tínhamos construído para nós nos seria, no final, vedado. Provavelmente ao som de uma voz que, polidamente, nos responderia aos nossos clamares com um lapidar "lamento, humanos, mas receio não poder fazer isso", como dizia o HAL-9000, na sua melhor impressão de um burocrata.

"(...) a ilustre carioca esboçou, na sua crónica, um sentimento de 2024: um espanto mudo com o infinito potencial duma inteligência que ultrapassa a nossa, seguido da realização de que o Shangri-La que tínhamos construído para nós nos seria, no final, vedado."

Bem se percebe a angústia; um combate frente a frente com uma inteligência artificial não augura nada que não uma sumária sova. Se houvesse um tale of the tape, seria algo como: "ChatGPT: numa tarde "leu" 40.000 artigos científicos, tendo redigido depois 200 páginas de manualística. Bento: numa tarde enviou 3 e-mails e leu 10 páginas da Teoria Geral do Direito Civil antes de passar 3 horas a ver tik-toks de piscinas a serem construídas no meio da selva." Comparação avassaladora, não só porque deixa a impressão de que o que nos define a nós, humanos, é a procrastinação, como também porque leva à conclusão de que mesmo que assim não fosse, não ganharíamos o frente-a-frente. Bento poderia ser um prodígio da leitura e da escrita, imune ao aliciamento do fluxo constante de conteúdo, e mesmo assim seria deixado a comer pó numa corrida de memória e conhecimento bruto com uma IA. Tentar saber ou processar mais informação do que uma IA para isso vocacionada é uma tarefa demasiado quixotesca até para o fulgor delirante do malfadado cavaleiro andante.

"Bem se percebe a angústia; um combate frente a frente com uma inteligência artificial não augura nada que não uma sumária sova".

Àqueles/as que não aproveitaram esta deixa para fugir para o Youtube sem sentimento de culpa, digo: não é uma batalha assim tão iníqua quanto parece, em primeiro lugar porque não tem de ser uma batalha de todo. Ao contrário dos moinhos quixotescos, ela é uma criação que veio da nossa consciência e engenho para nos ajudar a responder aos desafios do futuro, e não necessariamente uma antagonista. Ainda que seja autónoma, não é independente, e nessa medida o benefício para o Humano é o fundamento, linha orientadora, e limite da conduta destas IAs. Claro, circunstâncias

políticas, sociais e económicas poderão inviabilizar esse objetivo na prática, mas isso aplica-se a toda a inovação tecnológica, e não decorre de uma inevitabilidade fatal da IA.

Daqui resulta que entrar numa competição pessoal com uma IA não é só fútil, é desnecessário. Onerados desde o início dos tempos com o esforço e trabalho que até ora sempre foram a contrapartida do sustento e crescimento humanos, a chama que Prometeu nos trouxe dos Deuses liberta-nos, em parte, dessas grilhetas. Podemos enfim concentrar-nos um pouco menos no que o Prof. John Keating descreveu como "nobres demandas, necessárias ao sustento da vida", e um pouco mais nas "coisas pelas quais ficamos vivos". Para quê memorizar mais dígitos do pi, quando podemos descobrir como as cores funcionam numa tela? Podemos entregar a visão do Sr. Ford à IA, que nós retomamos o espírito pitagórico para fraccionar sons ao longo de uma corda. Sem as balizas da contingência humana a limitá-las, podemos tecer as palavras da forma que mais aprouver a cada um.

"Daqui resulta que entrar numa competição pessoal com uma IA não é só fútil, é desnecessário."

Dir-me-ão que já há IAs para fazer tudo isso, e que até esse reduto não passa de um ponto de passagem para o progresso delas. Qualquer pessoa pode criar paisagens infinitas e textos para os mais diversos fins com algumas palavras, e até os Beatles foram procurar a ajuda da IA para lançar o que provavelmente será o seu último single. É verdade. Porém, não haveria delírios da máquina sem séculos de devaneio humano, não haveria textos coligidos maquinalmente sem o esforço "manual" de gerações de artistas das letras, e sem a visão, engenho e nostalgia do John, do Paul e do George (vá, do Ringo também), Now And Then seria apenas um pastiche cansado de uma saudade fugidia. Sem os fundamentos humanos para a estribar, a IA estaria apenas a gesticular no ar.

"Qualquer pessoa pode criar paisagens infinitas e textos para os mais diversos fins com algumas palavras, e até os Beatles foram procurar a ajuda da IA para lançar o que provavelmente será o seu último single".

Na verdade, também é assim com o pensar e com o saber. Também a "carga pesada" cognitiva nos é tirada dos ombros, restando-nos a força para outras empresas. Já temos quem nos desenhe o mapa da ilhas do saber – só nos resta encontrar os caminhos entre elas. Já temos um manancial de informação, resta-nos aprender a separar criticamente o trigo do joio. Fomos, até agora, uma espécie de mineiros, no escuro a escavar por pepitas de conhecimento; hoje, já podemos ser uma espécie de ourives, encarregues de trabalhar a matéria-prima que nos é trazida para a mesa. Desonerados do labor que é começar a pensar, podemos agora levar esse pensamento mais longe e mais rápido.

D.

“Fomos, até agora, uma espécie de mineiros, no escuro a escavar por pepitas de conhecimento; hoje, já podemos ser uma espécie de ourives, encarregues de trabalhar a matéria-prima que nos é trazida para a mesa.”

“Deus criou um problema para si próprio. Ele terminará destruindo a máquina e recomeçando pela ignorância do homem diante da maçã. Ou o homem será um triste antepassado da máquina; melhor o mistério do paraíso.”, assim termina Lispector, com um vaticínio de quem conhece a humanidade. Não sei, Clarice. É verdade que estamos condenados a perder uma competição com a máquina, mas se usarmos a inteligência que temos e a juntarmos com a artificial, se calhar o paraíso fica um pouco menos misterioso.

“É verdade que estamos condenados a perder uma competição com a máquina, mas se usarmos a inteligência que temos e a juntarmos com a artificial, se calhar o paraíso fica um pouco menos misterioso.”

Miguel Mota

Assistente Convidado da Faculdade de Direito
(Católica de Lisboa)

D.

A MENTE É UM ÓTIMO SERVO, MAS UM PÉSSIMO DOMINUS

RECEIOS SOBRE A APLICAÇÃO DO DIREITO E O DESEJO DE A REPLICAR ATRAVÉS DE SISTEMAS DE INTELIGÊNCIA ARTIFICIAL

Em 2011, a revista *Nature* publicou um artigo que, infelizmente, passou praticamente despercebido pela comunidade jurídica. Alguns cientistas das Universidades Ben Gurion e Columbia examinaram mais de 1000 decisões de 8 juízes ao longo de um ano, visando descobrir os fatores que mais contribuíam para a decisão de admitir que um arguido saísse (ou não) em liberdade condicional. As conclusões do estudo foram arrebatadoras. Aparentemente, critérios jurídicos como a gravidade do crime cometido, ou a existência de precedentes criminais, pouco ou nada contribuíam para a decisão, uma vez que o critério que melhor explicava as decisões dos juízes era o seu… nível de açúcar no sangue. No início do dia, a probabilidade de admitir que um arguido saísse em liberdade condicional encontrava-se perto dos 65%, reduzia linearmente com o passar das horas até à hora do almoço – onde se encontrava perto dos 0% – e, depois da refeição, regressava aos 65%, voltando a reduzir até ao final do dia (*V.,* CORBYN, Zoë, Hungry judges dispense rough justice, Nature, 11/04/2011).

D.

"No início do dia, a probabilidade de admitir que um arguido saísse em liberdade condicional encontrava-se perto dos 65%, reduzia linearmente com o passar das horas até à hora do almoço (...)"

Estas desconfortáveis conclusões são corroboradas por outros estudos célebres. *Colorandi* causa, ao se analisarem as decisões prolatadas pelos tribunais juvenis do Louisiana entre 1996 e 2012, descobriu-se uma relação poderosa entre quão desfavoráveis as decisões judiciais eram para os arguidos e a existência de resultados inesperados em jogos de futebol americano. Derrotas inesperadas das equipas locais aumentavam as penas aplicadas na semana que se seguia ao resultado, mas vitórias inesperadas (ou derrotas contra equipas do mesmo nível) não determinavam qualquer efeito nas penas. (V. EREN, Ozkan, MOCAN, Naci, Emotional Judges and Unlucky Juveniles, *American Economic Journal: Applied Economics*, vol. 10, n. º 3, 2018, p. 171-205).

"(...) descobriu-se uma relação poderosa entre quão desfavoráveis as decisões judiciais eram para os arguidos e a existência de resultados inesperados em jogos de futebol americano."

Os estudos supra exigem reflexões profundas sobre o modo como o Direito é pensado e aplicado. Qualquer pretensão de cientificidade do Direito, que passe pelo controlo racional das suas conclusões tem, necessariamente, de ser sensível não apenas à heurística das suas conceções e sistemática, mas também à empírica dos resultados e método que afirma empregar. Contudo, os casos *supra* são particularmente delicados desde logo porque nunca um juiz atribuiria certa conclusão jurídica à sua fome ou ao facto da sua equipa de futebol ter perdido um jogo importante. Nestes casos, o viés associado à decisão não só é invisível ao próprio aplicador do Direito como, mais grave ainda, não é intuitivamente identificável enquanto tal pela população em geral. Isto significa, inexoravelmente, que dificilmente o legislador conseguirá evitar este tipo de resultados.

"Qualquer pretensão de cientificidade do Direito, que passe pelo controlo racional das suas conclusões tem, necessariamente, de ser sensível não apenas à heurística das suas conceções e sistemática, mas também à empírica dos resultados e método que afirma empregar."

Face a estas perplexidades, a definição de um qualquer sistema de inteligência artificial que vise replicar a prática jurídica convoca uma cautela extraordinária. A utilização destes programas pelo poder judicial é, ainda, vista com considerável ceticismo, mas este receio esmorece quando o objetivo é incorporá-los na prática administrativa – na verdade, o seu estudo começa a ser cada vez mais comum (por todos, TRALHÃO, Mariana, A utilização de inteligência artificial no processo decisório administrativo e os princípios da administração eletrónica, *Revista Jurídica AAFDL*, n.º

32/33, 2021, pp. 248 e ss.). No domínio administrativo, admite-se que as vantagens de interesse público que estes sistemas permitem obter ultrapassam os riscos associados (*V.* o Relatório elaborado pelo *Committee on Standards in Public Life*, para o Governo do Reino Unido, *Artificial Intelligence and Public Standards*, Fevereiro de 2020, pp. 12 e ss. e 57 e ss.). O meu objetivo, neste pequeno texto, foi, portanto, chamar à atenção para um problema que, de regra, passa despercebido aos juristas. É que a utilização destes sistemas não é apenas apta a gerar decisões enviesadas por racismo ou sexismo (*V.* GARCIA, Megan, Racist in the Machine, *World Policy Journal*, vol. 33, n. º 4, 2016/2017, p. 111-117), mas, pelo contrário, põe à prova a legitimidade de *todo* o sistema jurídico, *maxime* a correspondência entre os valores que o sistema jurídico proclama defender e a forma como efetivamente os defende.

É que a utilização destes sistemas não é apenas apta a gerar decisões enviesadas por racismo ou sexismo (...), mas, pelo contrário, põe à prova a legitimidade de *todo* o sistema jurídico, *maxime* a correspondência entre os valores que o sistema jurídico proclama defender e a forma como efetivamente os defende.

Gonçalo Sá Gomes

Aluno do Mestrado de Administrativo da Faculdade de Direito (Católica de Lisboa)

UM NOVO REGULADOR PARA A INTELIGÊNCIA ARTIFICIAL?

É já um lugar-comum afirmar que o desenvolvimento da Inteligência Artificial ("IA") tem suscitado – e irá, cada vez mais, continuar a suscitar – múltiplos problemas à Humanidade, em diversos campos. Um deles é, como não poderia deixar de ser, o campo do Direito, que não pode deixar de regular uma realidade que, porém, evolui de forma demasiado rápida para poder ser devidamente enquadrada por normas rígidas e estáticas. Há, pois, uma tensão inevitável entre o Direito e a IA, numa relação que não pode deixar de existir, mas que, ao mesmo tempo, nunca será pacífica: como já foi escrito, este desenvolvimento tecnológico que se processa de forma desconcentrada e acelerada desafia a evolução própria do Direito, que assenta no empirismo do saber e em modelos pré-estabelecidos (Cf. João Taborda da Gama, "*Inteligência Artificial e Fiscalidade*", *Inteligência Artificial & Direito*, Almedina, Coimbra, 2020, p. 233).

Neste cenário "sem rede" e sempre nas fronteiras da regulação, a primeira preocupação do Direito é a de garantir que o progresso tecnológico se faz com respeito pelos princípios éticos. Neste sentido, um pouco por todo o mundo, os Estados, enquanto Estados de Direito, começam a adotar regras nacionais para a assegurar que a IA é desenvolvida em conformidade com os direitos fundamentais dos cidadãos.

D.

"Neste cenário "sem rede" e sempre nas fronteiras da regulação, a primeira preocupação do Direito é a de garantir que o progresso tecnológico se faz com respeito pelos princípios éticos."

E aqui entra um novo desafio: se, por um lado, cada Estado não pode deixar de regular a utilização da IA, no legítimo exercício dos seus poderes de soberania, tendo em conta o seu contexto (social, económico, cultural, tecnológico, etc.) e de acordo com as suas próprias conceções e mundividência, por outro lado, não é desejável que a regulação da IA seja disciplinada de modo disperso, avulso e casuístico, com cada ordenamento nacional a adotar soluções específicas que não têm minimamente em consideração os modelos vigentes nos demais países e no espaço global em que todos se inserem. Em particular, no âmbito europeu, as instituições da União Europeia pretendem evitar que as necessárias regulações nacionais conduzam à fragmentação do mercado interno e reduzam a segurança jurídica para os operadores que desenvolvem, importam ou utilizam sistemas de IA.

"E aqui entra um novo desafio: se, por um lado, cada Estado não pode deixar de regular a utilização da IA, no legítimo exercício dos seus poderes de soberania, tendo em conta o seu contexto (...) e de acordo com as suas próprias conceções e mundividência, por outro lado, não é desejável que a regulação da IA seja disciplinada de modo disperso, avulso e casuístico, (...)."

É com este intuito que surge o Regulamento Europeu para a Inteligência Artificial ("Regulamento"), que visa estabelecer regras harmonizadas para garantir que os sistemas desenvolvidos e utilizados na Europa respeitam os valores e direitos da União, incluindo a privacidade, a igualdade, a segurança e o bem-estar social e ambiental, consoante o respetivo grau de risco.

Ora, um dos aspetos em que o Regulamento impõe uma uniformização aos vários Estados-membros é o da designação de uma "autoridade de supervisão do mercado" para acompanhar o cumprimento das suas disposições e para servir como ponto de contacto junto da Comissão. Ou seja, o Regulamento vem prever que exista uma entidade reguladora para a IA em cada um dos Estados-Membros da União Europeia.

Assim, a par de múltiplos outros pontos que, mais tarde ou mais cedo, terão de ser densificados ao nível nacional, a criação de um regulador para a IA surge como uma obrigação para o legislador português. O que justifica que se comece desde já o debate sobre o "figurino" deste futuro regulador, seja no que diz respeito às suas competências, seja, desde logo, à decisão entre a criação ex novo de um regulador para este efeito ou, diversamente, a atribuição destas funções a uma das entidades reguladoras já atualmente existentes em Portugal. E esse debate é incontornável, na medida em que o Regulamento não parece impor uma solução única, pelo que caberá a cada Estado-membro decidir sobre o concreto modelo a adotar.

D.

"(...) a criação de um regulador para a IA surge como uma obrigação para o legislador português."

Procurando dar o "pontapé de saída" para essa necessária reflexão, crê-se, numa primeira leitura, que ambos os modelos são exequíveis, apresentando, entre eles, diferentes vantagens.

Assim, a atribuição das competências previstas no Regulamento a uma entidade reguladora já existente poderá potenciar:

(i) **a celeridade e economia de meios**, porque se privilegiam os recursos já existentes e se alocam novos investimentos na criação de novo conhecimento especializado ao capital humano existente;
(ii) **o aproveitamento do know-how**, porque uma entidade reguladora estabelecida acumula experiência no setor em que atua, promovendo a eficiência na supervisão da utilização de sistemas de IA nesse setor, o que, em alguns casos, é já uma realidade; e
(iii) **a consistência e previsibilidade das decisões**, porque uma entidade reguladora já estabelecida tem um precedente decisório consistente, contribuindo para uma maior previsibilidade do funcionamento do mercado.

Por seu turno, as vantagens que podem apontar-se à criação de uma nova entidade reguladora especificamente para a IA passam pela:

(iv) **flexibilidade e especialização**, na medida em que que esta nova entidade reguladora poderia ser desenhada "à medida" e a sua estruturação poderia ser mais facilmente adaptável à evolução da tecnologia, focando-se nos desafios regulatórios próprios da IA;
(v) **independência**, na medida em que uma nova entidade estaria menos permeável a influências face aos regulados (por não atuar num único setor económico, antes assumir um âmbito horizontal ou transversal a todos os setores); e
(vi) **inovação**, na medida em que permitiria desenvolver, desde logo, uma nova dinâmica competitiva na área da IA com a criação de um novo quadro regulamentar de fomento ao investimento.

Como termo imediato de comparação, deve notar-se que aqui ao lado, em Espanha, a opção passou pela criação de uma nova entidade reguladora autónoma para a IA: com efeito, a recentíssima Lei n.º 22/2021, de 28 de dezembro, criou a Agência Espanhola de Supervisão da IA ("AESIA"), que será a entidade responsável pela supervisão em matéria de desenvolvimento e utilização de sistemas de IA, competindo à AESIA eliminar ou reduzir os riscos para a integridade, a

privacidade, a igualdade de tratamento e a não discriminação, e para outros direitos fundamentais que possam ser afetados pela utilização indevida dos sistemas de IA.

"Como termo imediato de comparação, deve notar-se que aqui ao lado, em Espanha, a opção passou pela criação de uma nova entidade reguladora autónoma para a IA (...)."

Em Portugal, tanto quanto se sabe, alguns reguladores já sinalizaram a sua disponibilidade para, se for essa a opção nacional, vir a assumir as funções previstas no Regulamento: é o caso, pelo menos, da ANACOM, que, segundo foi noticiado, se mostrou disponível para, "*dentro de toda esta transformação digital (...) contribuir para essa regulação e depois supervisionar todo o trabalho que é feito nessa área, de forma a proteger os cidadãos a proteger os países e evitar os riscos que, de facto, a inteligência artificial, nalguns casos, também pode potenciar*" (Cf. a entrevista do (então) presidente do Conselho de Administração da Autoridade Nacional de Comunicações à Agência Lusa, em 14 de dezembro de 2023. De notar, em qualquer caso, que esta entidade reguladora tem agora uma nova presidente, que iniciou funções há muito pouco tempo).

"Em Portugal, tanto quanto se sabe, alguns reguladores já sinalizaram a sua disponibilidade para, se for essa a opção nacional, vir a assumir as funções previstas no Regulamento: é o caso, pelo menos, da ANACOM (...)."

Não obstante, o debate ainda vai no início, não sendo possível antecipar qual a opção que irá ser tomada a este respeito pelo legislador, até porque, neste momento, se desconhece a composição partidária do Parlamento e a orientação política do futuro Governo que virá a ser nomeado na sequência das eleições legislativas do próximo dia 10 de março. Uma coisa, porém, é certa: esta escolha – criação de um novo regulador para a IA ou integração das suas competências num regulador já existente – não poderá deixar de ser feita na legislatura que está prestes a iniciar-se, pelo que será bom que os responsáveis políticos decidam de modo informado.

Marco Caldeira e Gonçalo Mesquita Ferreira

Associado Coordenador da VdA
Associado da VdA e alumnus da Faculdade de Direito (Católica em Lisboa)

D.

D.

O QUE PODEM ESPERAR DO REGULAMENTO INTELIGÊNCIA ARTIFICIAL OS EMPREGADORES QUE PROJETEM USAR SISTEMAS DE IA EM CONTEXTO LABORAL

POR HELENA TAPP BARROSO

Depois do acordo político atingido, em dezembro de 2023, sobre a proposta, apresentada pela Comissão Europeia, em 21 de abril de 2021, de regulamento destinado a estabelecer regras harmonizadas em matéria de inteligência artificial (IA) – e concluída a última reunião do trílogo no âmbito do processo legislativo ordinário da UE -, foi conhecido o texto quase final da peça transformadora da legislação europeia que será o Regulamento IA.

Este mês de fevereiro trouxe os marcos importantes da aprovação do Comité de Representantes Permanentes dos Governos dos Estados-Membros seguida da aprovação nas Comissões LIBE e IMCO do Parlamento Europeu. A votação no plenário do Parlamento Europeu está agendada para abril.

É uma peça legislativa que pretende alcançar o duplo objetivo de promover a adoção da IA e de abordar os riscos associados a determinadas utilizações desta tecnologia. Enquanto tal, inclui um leque vasto de obrigações e exigências de governação aplicáveis em vários pontos da cadeia de valor digital da IA, incluindo fornecedores e utilizadores de sistemas de IA.

As empresas, incontornavelmente, terão de rever e ajustar normas, políticas e processos, para preparar a conformidade futura com as regras que já podem ser antevistas.

As empresas, incontornavelmente, terão de rever e ajustar normas, políticas e processos, para preparar a conformidade futura com as regras que já podem ser antevistas.

Hoje, a realidade do uso de sistemas de IA em contexto laboral é já incontornável, prevendo-se o aumento do número dos utilizadores e uma amplificação e diversificação dos planos desse uso, seja no acesso à prestação de trabalho, na gestão dessa prestação ou nas vicissitudes dos vínculos que suportam a sua prestação.

D.

Nesse plano, os riscos potenciais mais relevantes, são riscos para os direitos fundamentais à proteção de dados e à privacidade de trabalhadores – especialmente nos sistemas de IA utilizados para monitorizar o desempenho e o comportamento – e o risco da discriminação, designadamente, mas não exclusivamente, com base no género, pois os sistemas de IA podem perpetuar padrões históricos de discriminação (*algorithmic bias*), como já registado em diversos casos conhecidos.

"(...) os riscos potenciais mais relevantes, são riscos para os direitos fundamentais à proteção de dados e à privacidade de trabalhadores (...) e o risco da discriminação (...)."

Há riscos associados à opacidade do processo de tomada de decisões. A explicabilidade dos *outputs* deve garantir-se, evitando aplicações de IA, conhecidas como, *black box*. A transparência, quanto aos dados usados e ao modo de funcionamento (não técnico) do modelo, são essenciais para garantir a confiança no seu uso e evitar potenciais disfunções.

As alterações da Lei n.º 13/2023, de 3 de abril de 2023 ao Código do Trabalho (CT) introduziram disposições respeitantes ao uso de IA na tomada de decisões.

Dessas, muito sumariamente, destacamos:

- a previsão de deveres de prestação de informação a favor de trabalhadores e respetivas estruturas representativas, sobre "os parâmetros, os critérios, as regras e as instruções em que se baseiam os algoritmos ou outros sistemas de inteligência artificial que afetam a tomada de decisões sobre o acesso e a manutenção do emprego, assim como as condições de trabalho, incluindo a elaboração de perfis e o controlo da atividade profissional – artigo 106.º, n.º 2 al. s) (trabalhadores), artigo 424.º, n.º 1, al. f) (comissão de trabalhadores) e 466.º, n.º 1 al. d) (delegados sindicais), prevendo-se neste último caso também a consulta –; e
- a alusão expressa de que os direitos à igualdade de oportunidades e de tratamento no acesso ao emprego, na formação e promoção ou carreira profissionais e nas condições de trabalho, não podem ser prejudicados pela "tomada de decisões baseadas em algoritmos ou outros sistemas de inteligência artificial" – artigo 24.º, n.º 3 do CT –.

O legislador laboral reporta-se, segundo a sua própria expressão, ao uso de "algoritmos ou outros sistemas de inteligência artificial", expressão que carece de rigor, na parte em que, parece, identifica os processos de decisão algorítmica com a IA. Estas normas têm, por isso, nesta referência ao uso de algoritmos, um âmbito de aplicação que não coincide, integralmente, com o do Regulamento de IA.

D.

No quadro do Regulamento de IA são sistemas de IA os sistemas *machine-based* concebidos para funcionar com diversos níveis de autonomia e que pode apresentar adaptabilidade depois de implementado e que, para objetivos explícitos ou implícitos, infere, a partir dos inputs recebidos, como gerar resultados tais como previsões, conteúdos, recomendações ou decisões, que podem influenciar ambientes físicos ou virtuais. O Regulamento não pretende abranger os sistemas de software tradicionais mais simples ou abordagens de programação, que se baseiam em regras definidas exclusivamente por pessoas singulares para executar operações automaticamente.

Vejamos, no quadro do Regulamento de IA – e, por referência à redação pré-final conhecida –, as regras a ter em conta no uso, por empregadores, de sistemas de IA em contexto laboral.

O aspeto mais relevante prende-se com os chamados sistemas de IA de risco elevado. Contemplam-se regras de classificação de sistemas de IA como de *risco elevado*, associando-se-lhe um regime próprio. Prevê-se ainda uma listagem de sistemas de *risco elevado*, em função da finalidade prevista, e do nível de risco de danos para a saúde e a segurança ou de prejuízo para os direitos fundamentais das pessoas, que representam, tendo em conta a sua gravidade e probabilidade de ocorrência, quando utilizados num conjunto de domínios predefinidos.

"O aspeto mais relevante prende-se com os chamados sistemas de IA de risco elevado."

Aquela listagem contempla o uso de sistemas de IA no plano do emprego, gestão de trabalhadores e acesso ao emprego por conta própria.

Concretamente, são classificados como sistemas de IA de risco elevado, os concebidos para serem utilizados:

a) no recrutamento ou na seleção de pessoas singulares, designadamente, para divulgação seletiva de anúncios de emprego, para análise ou filtragem de candidaturas e para avaliação de candidatos; e

b) para tomada de decisões que afetem as condições de relações de prestação de trabalho, promoções ou cessações de relações contratuais de trabalho, para repartição de tarefas baseadas em comportamentos individuais ou em características ou traços pessoais, e no controlo e avaliação do desempenho e do comportamento de pessoas envolvidas nessas relações.
Diferentemente do que sucede com as normas do CT mencionadas supra, os sistemas listados não se reportam apenas ao trabalho subordinado, abrangendo também a realidade do "acesso ao trabalho por conta própria".

D.

Os sistemas de IA de risco elevado ficarão sujeitos a exigências respeitantes à qualidade dos conjuntos de dados utilizados, a documentação técnica e manutenção de registos, regras de transparência e de prestação de informações aos utilizadores, requisitos de supervisão humana, solidez técnica relativamente aos riscos associados às suas limitações (como erros e incoerências) e requisitos de exatidão e cibersegurança, de acordo com o estado da técnica, tudo aspetos destinados à mitigação eficaz dos riscos.

Há deveres aplicáveis aos fornecedores, enquanto entidades que desenvolvem o sistema ou que têm o sistema desenvolvido e que o colocam no mercado ou em serviço sob o seu nome ou marca. Entre eles, incluem-se os deveres de estabelecer um sistema de gestão da qualidade sólido, de garantir a realização de um procedimento de avaliação de conformidade, de elaborar a documentação pertinente e de estabelecer um sistema de acompanhamento pós-comercialização capaz.

O Regulamento IA também prevê deveres aplicáveis aos utilizadores de sistemas de IA de risco elevado.

O utilizador do sistema de IA é a entidade que utiliza, sob a sua autoridade, um sistema de IA.

O utilizador do sistema de IA é a entidade que utiliza, sob a sua autoridade, um sistema de IA.

Quando ocorrer uso de sistemas, pelo empregador, potencial empregador ou entidade contratante, no recrutamento ou seleção, na tomada de decisões que afetem as condições para a prestação do trabalho, na repartição de tarefas em função de comportamentos individuais ou características ou traços pessoais, nas decisões para promoções, no controlo e na avaliação do desempenho e do comportamento dos que prestam trabalho, e ainda, na tomada de decisões para a cessação de relações de trabalho, o empregador assume a qualidade de utilizador do sistema de IA de risco elevado, ficando, nessa medida, sujeito aos correspondentes deveres.

"(...) o empregador assume a qualidade de utilizador do sistema de IA de risco elevado, ficando, nessa medida, sujeito aos correspondentes deveres."

Estes deveres incluem os de adoção de medidas técnicas e organizativas adequadas para garantir que o uso dos sistemas seja conforme com as instruções de utilização, particularmente, na aplicação de medidas de supervisão humana indicadas pelo fornecedor.

Há também deveres respeitantes à conservação dos *logs* gerados automaticamente pelo sistema (sempre que esses *logs* estejam sob o controlo do utilizador), a assegurar que os dados de entrada dos sistemas sejam pertinentes e suficientemente representativos, em função da sua finalidade (na medida em que o utilizador exerça controlo sobre esses dados).

D.

Os utilizadores de sistemas de IA de risco elevado também ficam sujeitos a deveres de monitorização do funcionamento dos sistemas, deveres de comunicação ao fornecedor ou outros operadores e também à autoridade de fiscalização (e.g. de riscos identificados na sua utilização ou de determinadas ocorrências), estando também previstos alguns deveres de suspensão da utilização do sistema.

Os deveres exemplificados *supra*, são universalmente aplicáveis aos utilizadores de sistemas de IA de risco elevado e, nessa medida, também aos empregadores que os utilizem em contexto laboral. A estes, acrescem outros deveres – de informação e transparência –, nos dois casos seguintes: (i) uso de sistemas de IA de risco elevado para assistir em processos de tomada de decisões respeitantes a pessoas singulares, e (ii) no uso que os empregadores façam de sistemas de IA de risco elevado em contexto laboral.

"Os deveres exemplificados supra, são universalmente aplicáveis aos utilizadores de sistemas de IA de risco elevado e, nessa medida, também aos empregadores que os utilizem em contexto laboral."

Quanto ao primeiro, os utilizadores deverão informar as pessoas a quem as decisões respeitem, de que recorrem ao sistema de IA de risco elevado, na tomada de decisões. Quando o empregador (ou potencial empregador), use sistemas de IA de risco elevado para o assistir em processos de tomada de decisões respeitantes a trabalhador (ou candidato), esse dever poderá ser-lhe aplicável.

Entre este dever e o dever de prestar informação ao trabalhador previsto no CT, há uma zona de intersecção, na parte em que o CT respeita à utilização de sistemas "(…) que afetam a tomada de decisões sobre o acesso e a manutenção do emprego, assim como as condições de trabalho (…)" mas o CT densifica a informação a prestar. Neste caso, ela versa, concretamente, sobre "os parâmetros, os critérios, as regras e as instruções em que se baseiam os (…) sistemas de inteligência artificial". Estes deveres não excluem outros deveres de informação, como os decorrentes do RGPD, muito em especial – noutro plano de intersecção – nas aplicáveis em caso de tomada de decisões (exclusivamente) automatizadas (cf. artigos 13.º, n.º 2 al. f), 14.º, n.º 2, al. f) e artigo 22.º do RGPD).

Quanto ao segundo, o regulamento prevê que sempre que planeie adotar o uso de um sistema de IA de risco elevado em contexto laboral, o empregador deve prestar informação sobre o facto de os trabalhadores ficarem abrangidos por esse uso, antes de iniciar esse mesmo uso.

D.

A informação é prestada aos trabalhadores afetados e, também, às respetivas estruturas representativas, prevendo-se que isso seja feito, em conformidade com as regras e procedimentos previstos na legislação nacional sobre a prestação de informação aos trabalhadores e respetivas estruturas representativas. Neste ponto, poder-se-ia antever outra área de intersecção, entre deveres. Porém, se atentarmos no objeto da informação a prestar e, mesmo, no momento do seu cumprimento, verifica-se que não são coincidentes, ainda que em ambos os casos, a transparência sirva a mitigação do mesmo risco.

A aplicação de diferentes conjuntos de normas do Regulamento de IA ocorrerá em momentos distintos no tempo.

Uma vez publicado no jornal oficial da EU – e identificada a data da sua entrada em vigor que ocorrerá no vigésimo dia seguinte ao da publicação – as obrigações respeitantes aos sistemas de IA de elevado risco aplicam-se decorridos vinte e quatro meses da entrada em vigor. É a partir desse momento que os empregadores, enquanto utilizadores dos sistemas de IA de risco elevado, ficam obrigados ao cumprimento dos correspondentes deveres.

Vale, ainda, a pena mencionar que o Regulamento de IA proíbe algumas práticas de IA, entre as quais, com relevância para o contexto da prestação de trabalho, a colocação no mercado, a colocação em serviço ou a utilização de sistemas de IA destinados a inferir emoções de pessoas singulares em contexto laboral. Estas proibições aplicam-se seis meses após a data de entrada em vigor do Regulamento de IA.

Como nota final, refira-se que o regulamento permite aos Estados-Membros ou à União manter ou introduzir disposições legislativas, regulamentares ou administrativas mais favoráveis aos trabalhadores, em matéria de proteção dos seus direitos, relativamente à utilização de sistemas de IA pelos empregadores, ou disposições que incentivem ou permitam a aplicação de convenções coletivas, contendo previsões, igualmente, mais favoráveis.

"(...) o regulamento permite aos Estados-Membros ou à União manter ou introduzir disposições legislativas, regulamentares ou administrativas mais favoráveis aos trabalhadores (...)."

Helena Tapp Barroso
Sócia da Morais Leitão

D.

DESAFIOS DA INTELIGÊNCIA ARTIFICIAL GENERATIVA NO QUADRO JURÍDICO DA UNIÃO EUROPEIA

UMA BREVE ANÁLISE

A ascensão da Inteligência Artificial Generativa (IAG), particularmente os Modelos de Linguagem Grandes (LLMs), como o ChatGPT, marca uma revolução na tecnologia de IA, apresentando novos desafios e oportunidades para o quadro jurídico da União Europeia (UE). Este artigo visa abordar alguns destes temas, não nos debruçando em particular sobre o Regulamento de Inteligência Artificial (RAI) mas sim em temas que gravitam à volta do Regulamento.

Estando atualmente a regulação da IA na UE centrada no RAI, a natureza inovadora e complexa da IAG requer uma visão holística sobre o tema, sendo necessário ter em linha de conta diferentes realidades, como a responsabilidade civil, a propriedade intelectual, privacidade e cibersegurança.

"Estando atualmente a regulação da IA na UE centrada no RAI, a natureza inovadora e complexa da IAG requer uma visão holística sobre o tema (...)."

As propostas legislativas sobre responsabilidade civil e IA, como a Diretiva sobre Responsabilidade por Produtos Defeituosos (PLD) e a Diretiva de Responsabilidade por Inteligência Artificial (AILD), vêm colmatar algumas lacunas existentes na legislação, especialmente no que diz respeito a estas tecnologias invadoras.

Na PLD, que se foca essencialmente na responsabilidade contratual e nos danos causados por produtos defeituosos, a definição de produto foi alterada de modo a abranger, por exemplo, as atualizações de software, a inteligência artificial (IA) ou os serviços digitais (tais como os robôs, os drones ou os sistemas de casas inteligentes), excluindo software livre ou de código aberto. Reconhece-se agora que um sistema de IA pode-se tornar defeituoso com base no conhecimento adquirido/aprendido após a sua implementação, estendendo a responsabilidade civil a essas situações.

D.

Já relativamente à AILD, a mesma visa a responsabilidade extracontratual causada por sistemas de IA, procurando harmonizar as regras europeias e garantir meios eficazes para identificar os responsáveis pelos danos e os elementos de prova relevantes para uma determinada ação indemnizatória. Estas regras aplicam-se independentemente de serem sistemas de IA de alto risco ou não, nos termos do RAI.

É importante frisar que ambas as diretivas mencionadas reconhecem a potencial opacidade da IA e o desequilíbrio de informações entre *developers* e utilizadores/ consumidores, introduzindo mecanismos de divulgação e presunções ilidíveis que colocam o ónus de prova em quem desenvolve este tipo de algoritmos.

"É importante frisar que ambas as diretivas mencionadas reconhecem a potencial opacidade da IA e o desequilíbrio de informações (...)."

Também o tema da Propriedade Intelectual (PI) é bastante relevante, na medida em que estes modelos, ao processarem e gerarem novos conteúdos a partir de vastas quantidades de dados, levantam questões sobre a autoria, originalidade e o uso legítimo de materiais protegidos por direitos de autor.

"(...) estes modelos, ao processarem e gerarem novos conteúdos a partir de vastas quantidades de dados, levantam questões sobre a autoria, originalidade e o uso legítimo de materiais protegidos por direitos de autor."

A formação de LLMs, que envolve o uso de grandes conjuntos de dados, incluindo materiais protegidos por direitos de autor, coloca em primeiro plano a questão da licitude da utilização de obras protegidas no treino do algoritmo. Neste contexto, a exceção *text* and *data mining* apresenta-se como uma solução potencial, mas a sua aplicabilidade é complexa e restringe-se a determinadas condições estabelecidas pelo quadro legal, destacando a necessidade de uma abordagem mais clara e abrangente que permita o treino do algoritmo sem violar direitos autorais.

Por outro lado, a geração de outputs por LLMs suscita questões sobre se esses outputs podem ser considerados criações autónomas e originais e, portanto, protegíveis sob as normas de PI. A distinção entre o uso de LLMs como ferramentas para potenciar a criatividade humana e a sua operação com um grau significativo de autonomia será, pois, fundamental neste tipo de análise. Enquanto que o uso assistido se pode manter sob o "guarda-chuva" da proteção de PI atualmente existente, a proteção dos outputs gerados de forma autónoma enfrenta alguns desafios legais, o que parece levar à necessária atualização do quadro legal existente de modo a que se reconheça a contribuição tecnológica sem subverter os fundamentos e princípios da PI.

"(...) a geração de outputs por LLMs suscita questões sobre se esses outputs podem ser considerados criações autónomas e originais e, portanto, protegíveis sob as normas de PI."

D.

Relativamente à cibersegurança, a complexidade e a alta dimensionalidade dos LLMs toram este tipo de tecnologia suscetível a ataques direcionados que procuram manipular os outputs gerados, torna necessário acautelar medidas técnicas robustas que permitam prevenir e controlar esses ataques. Estas questões já são, em certa medida, abordadas a nível legislativo, mas a eficácia na prevenção de ciberataques requer um o foco mais específico e adaptativo que vá além dos sistemas de alto risco identificados pelo RAI.

"Relativamente à cibersegurança, a complexidade e a alta dimensionalidade dos LLMs toram este tipo de tecnologia suscetível a ataques direcionados que procuram manipular os outputs gerados (...)."

Além disso, a capacidade dos LLMs de gerar "alucinações" — outputs plausíveis mas factualmente incorretos — sublinha a necessidade de abordagens inovadoras para validar a precisão dos dados gerados e minimizar a propagação de desinformação. Torna-se essencial estabelecer diferentes abordagens estratégicas, como por exemplo *Context Injection, Prompt Engineering, Retrieval-Augmented Generation* e *Domain- specific Fine-Tuning*.

Por fim, e no campo da privacidade e a proteção de dados, que também são áreas críticas na utilização deste tipo de tecnologia, dada a capacidade dos LLMs de processar enormes quantidades de informação que permita identificar pessoas singulares, importa referir que o Regulamento Geral sobre a Proteção de Dados (RGPD) oferece hoje um quadro robusto de proteção de alguns direitos fundamentais em matéria de privacidade, contudo o uso destas tecnologias exige algumas orientações específicas para assegurar a conformidade e proteger a privacidade dos indivíduos.

Em suma, podemos afirmar que a IAG representa uma área promissora de inovação, com o potencial de transformar diversos setores. No entanto, para maximizar os benefícios e minimizar os riscos, torna-se essencial que a UE adapte o seu quadro jurídico para enfrentar os desafios únicos apresentados por esta tecnologia. Através de uma abordagem regulatória equilibrada, não limitadora da inovação e numa base de abordagem ao risco, a UE tem agora ferramentas que permitirão um uso positivo da IAG, garantindo um equilíbrio entre os seus riscos e os seus benefícios.

Nuno Lima da Luz

Associado Sénior de TMT da Cuatrecasas

D.

PRECISÃO E PERCEÇÃO

COMO A IA ESTÁ REMODELANDO A DINÂMICA DE NEGOCIAÇÃO

Na arena dinâmica das negociações, ocorre uma transformação silenciosa, mas profunda, em que a intuição humana converge com a destreza analítica da Inteligência Artificial (IA). Esta exploração aventura-se na interação matizada de vantagens, considerações éticas e potenciais colaborativos que definem a ascensão provocadora da IA nas negociações.

A integração da IA nas negociações apresenta uma mudança de paradigma, oferecendo vantagens analíticas incomparáveis. Podemos imaginar negociações guiadas por algoritmos capazes de processar vastos conjuntos de dados a velocidades além da compreensão humana. Por exemplo, em fusões e aquisições, a IA torna-se um aliado estratégico ao avaliar rapidamente o sucesso potencial de uma operação societária. Ele disseca relatórios financeiros, condições de mercado e até mesmo perceções nas redes sociais, fornecendo aos negociadores insights sobre os meandros do negócio que teriam levado uma equipe de especialistas humanos consideravelmente mais tempo para analisar.

"A integração da IA nas negociações apresenta uma mudança de paradigma, oferecendo vantagens analíticas incomparáveis".

D.

No entanto, à medida que trilhamos este caminho transformador, as considerações éticas emergem como um ponto crítico de discórdia. A transparência dos algoritmos, potenciais enviesamentos incorporados nos dados de fonte e o reforço inadvertido das desigualdades existentes exigem um escrutínio rigoroso. Considere um cenário em que a IA é usada para avaliar as negociações salariais dentro de uma organização. Devem ser estabelecidos quadros éticos para garantir a equidade, a transparência e a responsabilização, evitando que enviesamentos inadvertidos afetem os resultados das negociações.

Ao contrário dos receios de que a IA substitua os negociadores humanos, o discurso introduz o conceito de aumento cognitivo. Aqui, a IA atua como colaboradora, libertando os humanos de tarefas analíticas rotineiras para se concentrarem nas dimensões relacional e estratégica das negociações. Nas negociações comerciais globais, a tradução de idiomas alimentada por IA facilita a melhoria da comunicação, quebrando barreiras linguísticas e permitindo que os negociadores se concentrem na construção de relacionamento e na compreensão das nuances culturais críticas para o sucesso de qualquer tomada de decisão.

As capacidades preditivas da IA, particularmente na antecipação de táticas de negociação, cativam o mundo académico. Ao discernir padrões nos comportamentos de negociação através de uma extensa análise de dados, a IA proporciona aos negociadores uma vantagem estratégica. Nas negociações em sede de conflito laboral, a IA pode prever potenciais pontos de conflito com base em dados históricos, permitindo que os negociadores abordem proativamente as preocupações e moldem estratégias que otimizem os resultados para todas as partes envolvidas.

“As capacidades preditivas da IA, particularmente na antecipação de táticas de negociação, cativam o mundo académico”.

Na busca de negociações justas e eficientes, a IA auxilia ainda mais com o conceito de licitação visual duplamente cega. Esta abordagem inovadora garante o anonimato no processo de submissão de propostas, reduzindo a influência de enviesamentos e promovendo condições de concorrência mais equitativas. Em negociações de contratação complexas, a licitação visual duplamente cega permite que as partes apresentem suas propostas sem revelar sua identidade, promovendo a transparência e a confiança. Esta aplicação da IA não só agiliza o processo de negociação, mas também aborda preocupações éticas, minimizando o impacto das afiliações pessoais no resultado.

À medida que o discurso se desenrola, tanto as comunidades académicas e profissionais mergulham nos desafios colocados pela incursão da IA nos meandros das interações humanas, das diferenças culturais e da inteligência emocional. Embora a IA se destaque em tarefas analíticas,

D.

navegar pelas sutilezas da dinâmica humana continua sendo um desafio contínuo. Por exemplo, a IA pode ter dificuldade em interpretar sinais não verbais ou compreender as subcorrentes emocionais numa sala de negociação, enfatizando a necessidade de negociadores humanos para complementar as capacidades da IA nestas áreas.

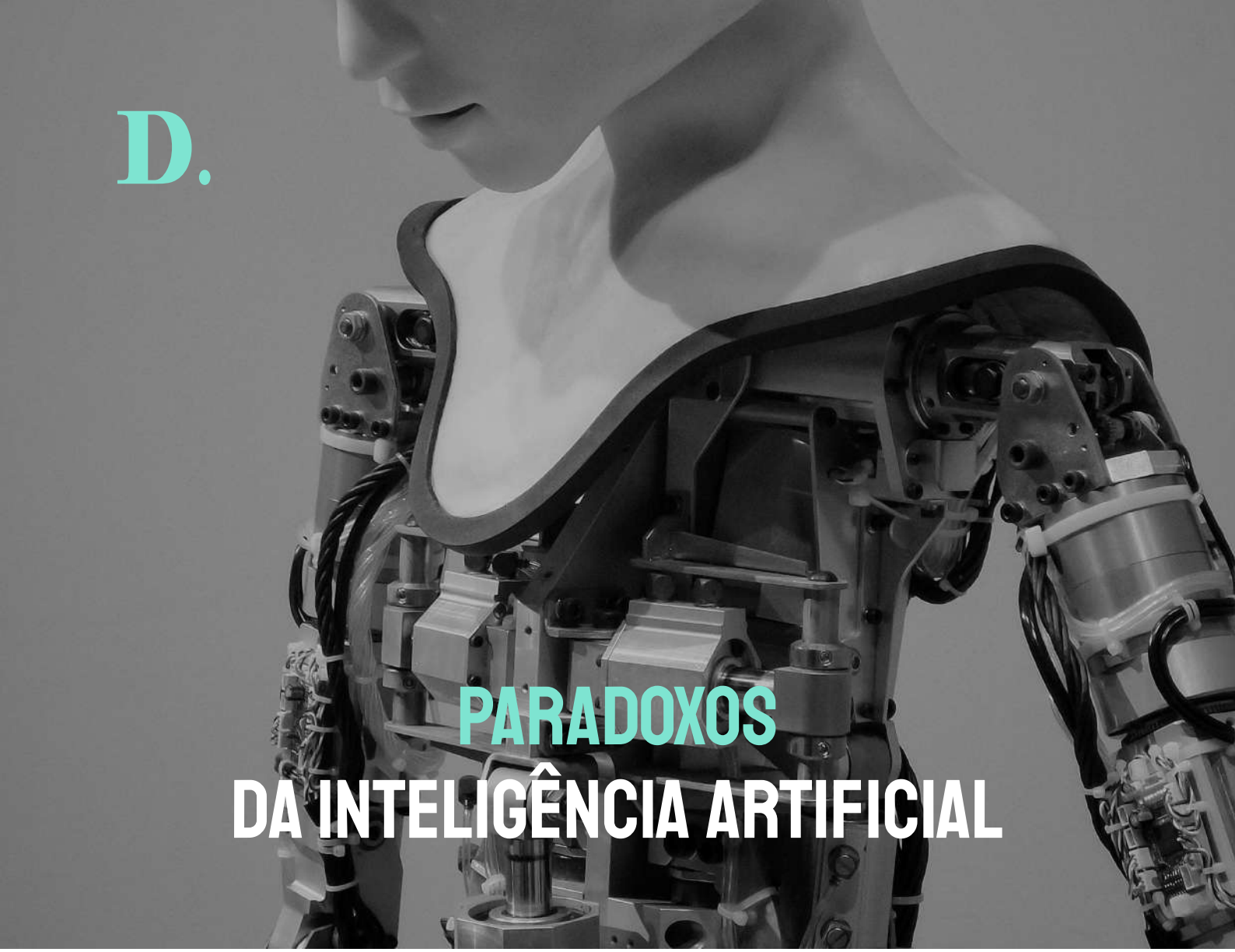
"Embora a IA se destaque em tarefas analíticas, navegar pelas sutilezas da dinâmica humana continua sendo um desafio contínuo".

Ao navegar no nexo da IA nas negociações, encontramo-nos na encruzilhada da promessa e da complexidade. A revolução silenciosa dos algoritmos entrelaçados com a intuição humana abre portas para vantagens ainda por imaginar. Considerações éticas, potenciais colaborativos e a essência insubstituível das habilidades humanas de negociação devem guiar o nosso caminho. À medida que abraçamos esta era transformadora, a verdadeira habilidade do negociador pode estar em orquestrar uma sinfonia onde a precisão analítica da IA, exemplificada em análise de M&A, negociações salariais, análise preditiva, por exemplo, se harmoniza com a subtileza e humanidade da nossa experiência em relações interpessoais.

Thomas Gaultier

Assistente Convidado da Faculdade de Direito (Católica do Porto)

D.



PARADOXOS DA INTELIGÊNCIA ARTIFICIAL

BIG DATA ANALYTICS E O "E" EM ESG

Inteligência Artificial ("IA") e Sustentabilidade medida por critérios Environmental Social Governance ("ESG") são dois dos temas da década. Natural que a interseção dos dois se tenha transformado numa discussão relevante.

Enquanto se discute a utilidade das ferramentas de IA para o cumprimento de métricas ESG, algumas vozes questionam se o custo daquelas não ultrapassa o seu benefício, às quais acrescentamos uma preocupação: a da eficiência dos resultados extraídos de algoritmos opacos.

Surge um duplo paradoxo: (i) o benefício de eficiência compensa o custo da utilização de IA? (ii) a IA é um mecanismo de auxílio eficaz para análise de métricas ESG?

D.

Por um lado, IA é a utilização de tecnologia para a automatização de tarefas que implicam inteligência humana. As instituições europeias, no mesmo sentido, indicam que "um sistema IA é um programa informático (...) capaz de, tendo em vista um determinado conjunto de objetivos definidos por seres humanos, criar resultados, tais como conteúdos, previsões, recomendações ou decisões, que influenciam os ambientes com os quais interage.". No âmbito desta análise, as ferramentas de IA que nos interessam são as de big data analytics, capazes de automatizar o armazenamento e análise de grandes volumes de dados (via machine learning), quanto a obtenção e extração de resultados.

"(...) IA é a utilização de tecnologia para a automatização de tarefas que implicam inteligência humana".

Por outro, ESG designa critérios para que as empresas forneçam informação acerca do seu impacto no mundo e do impacto que os riscos ambientais e sociais têm no seu desempenho (dupla materialidade). Mais concretamente, e embora os detalhes da informação possam variar de interesse para os vários stakeholders, importará que a empresa divulgue, por exemplo, quais as emissões de carbono que causa, como trata os seus trabalhadores e cuidado que tem na sua cadeia de valor e fornecimento e se tem a funcionar mecanismos de controlo da ética de decisões e, bem assim, se a remuneração dos administradores é desproporcionada em relação à média dos trabalhadores da empresa.

Podemos questionar: a IA ajuda no cumprimento de objetivos ESG? Na nossa opinião, sim, mas há riscos a gerir. Um dos grandes desafios que as empresas enfrentam na implementação e cumprimento de métricas ESG é na fase de obtenção e gestão de dados, pois provêm de fontes diversas, e.g., os dados de consumo energético, de água, papel, combustíveis fósseis, entre outros, não derivam da mesma fonte e muito menos estão necessariamente agrupados.

As ferramentas de IA permitem otimizar a obtenção e análise de conjuntos de dados a que dificilmente se chegaria numa análise humana documento por documento. Ou seja, a automatização traz benefícios à eficiência do cumprimento das métricas ESG, não só porque mecaniza a organização dos dados, como porque deles extrai recomendações de ação.

"As ferramentas de IA permitem otimizar a obtenção e análise de conjuntos de dados a que dificilmente se chegaria numa análise humana documento por documento."

D.

Acontece que alimentar e manter estas ferramentas implica um aumento de capacidade e conservação de um servidor físico, que tem uma pegada ecológica relevante. Por isso o primeiro paradoxo: a ferramenta que se quer utilizar para aumentar a observância de métricas ESG não obedece per se a essas métricas.

Contudo, ainda que haja custos para a manutenção, alimentação e análise automatizada inteligente, o benefício é superior: (i) há resultados relevantes que podemos extrair das decisões algorítmicas quanto ao comportamento ambiental da empresa; (ii) os próprios algoritmos são capazes de criar soluções para diminuir a sua própria pegada carbónica. O primeiro paradoxo parece ultrapassável.

Mas, um administrador, quando confrontado com a necessidade de tomada de decisões, necessita de ter à disposição várias opções, pois a tomada de uma decisão pode implicar não tomar outra. Para conseguir decidir se uma decisão é mais sustentável do que outra é necessário um racional alinhado com o propósito empresarial. Aliás, os responsáveis pelo cumprimento de métricas ESG estão incumbidos de fazer isso mesmo – e.g., explicar porque é melhor adotar uma no paper policy em vez de limitar o aluguer de veículos a carros elétricos devido à materialidade de cada tema.
O que sabemos é que os algoritmos que dão resultados mais corretos são os mais complexos, por seu turno, opacos quanto ao racional seguido até à recomendação. Surge o segundo paradoxo: o algoritmo justifica o investimento financeiro e ecológico necessário para me sugerir uma solução cuja lógica não compreendo? Como posso justificar que adotei uma decisão em detrimento de outra?

"(...) um administrador, quando confrontado com a necessidade de tomada de decisões, necessita de ter à disposição várias opções, pois a tomada de uma decisão pode implicar não tomar outra."

Para o resolver defendemos a adoção de algoritmos transparentes, trocando a complexidade por explicabilidade. Na realidade, não é impossível que através de algoritmos mais rudimentares se atinja o mesmo objetivo visado pelos modelos opacos.

O facto de inexistir precisão absoluta, salvaguardando a capacidade de compreensão de um racional, poderá ser benéfico. A menor precisão do algoritmo implica segundo controlo das equipas designadas para este efeito e do próprio administrador e nesse momento seriam escolhidas as medidas a adotar para maximizar cumprimento de métricas ESG, através do cruzamento entre resultado algorítmico, dados disponíveis e necessidades empresariais.

D.

Assim, os Paradoxos parecem ultrapassáveis, o primeiro in natura, o segundo garantindo que o algoritmo utilizado serve o seu propósito de auxílio na tomada de decisão de forma plena, e não apenas parcialmente.

Certo é que os temas analisados estão em evolução e encontramo-nos num momento chave para a sua intersecção: a sustentabilidade será uma preocupação do consumidor, obrigando as empresas a procurar soluções para atingir as métricas impostas (que por ora se chamam ESG); a IA, como facilitador de análise de dados, permite, desde que escolhido o algoritmo certo, uma análise de opção e tomada de decisão mais informada – não uma total delegação da decisão.

"Certo é que os temas analisados estão em evolução e encontramo-nos num momento chave para a sua intersecção (...)."

Já Fernando Pessoa escrevia em 1926 o que podemos ecoar, decorrido um século:
"A tendência moderna para a organização e coordenação, quanto possível perfeita, dos serviços de modo a torná-los mais simples e mais rápidos, deu em resultado a invenção, constantemente crescente, de sistemas, processos, móveis e aparelhos diversa e diferentemente conducentes a esse fim. Alguns desses processos, desses móveis e dessas máquinas são muito engenhosos; quase todos são úteis ou aproveitáveis. Mas o emprego deles, sejam quais forem, deve obedecer sempre a um critério superior. Um sistema não é uma cabeça; um móvel não é gente. Todos os processos e todos os aparelhos resultarão elementos inúteis de organização se as cabeças dos indivíduos que os empregam não estiverem organizadas também."

Maria Figueiredo e Lourenço Quintão

Of Counsel da CMS
Alumnus da Faculdade de Direito e Advogado Estagiário da CMS
(Católica no Porto)

D.

O SISTEMA FINANCEIRO COMO LABORATÓRIO DA INTELIGÊNCIA ARTIFICIAL

POR CLARA MARTINS PEREIRA

Em 14 de dezembro de 2021, a Oxford Union, um dos palcos de debate mais antigos do mundo Ocidental, testemunhou uma discussão histórica. Como de costume, a moção de debate acordada para essa noite—expressa ao estilo típico do Parlamento Britânico como "esta Casa acredita que a Inteligência Artificial nunca poderá ser ética"—foi objeto de discussão acesa, mas o debate contou com um protagonista particularmente invulgar: a própria Inteligência Artificial.

Convidado a pronunciar-se a favor da moção, o Megatron Transformer, uma forma de Inteligência Artificial criada pela NVIDIA (com base em trabalho desenvolvido pela Google), argumentou que é impossível criar uma Inteligência Artificial ética, e que a melhor forma de nos defendemos das consequências nefastas de uma Inteligência Artificial amoral seria eliminá-la (e eliminar o próprio Megatron) inteiramente. De seguida, o Megatron argumentou em sentido exatamente contrário.

D.

A título de experiência, o Megatron havia sido convidado a pronunciar-se sucessivamente a favor e contra a moção discutida na Oxford Union—acabando, na segunda parte do debate, a tecer uma história otimista sobre o papel da Inteligência Artificial na criação de um futuro mais risonho do que aquele ao alcance dos mais bem-intencionados humanos. Antes do debate, o Megatron tinha sido treinado longamente com a ajuda de múltiplos dados, incluindo a totalidade da Wikipédia e milhares de artigos noticiosos e "posts" do Reddit—permitindo-lhe esta aprendizagem argumentar convincentemente em favor de posições diametralmente opostas.

"A título de experiência, o Megatron havia sido convidado a pronunciar-se sucessivamente a favor e contra a moção discutida na Oxford Union—acabando, na segunda parte do debate, a tecer uma história otimista sobre o papel da Inteligência Artificial na criação de um futuro mais risonho do que aquele ao alcance dos mais bem-intencionados humanos."

Em última análise, o resultado deste debate é paradigmático das dificuldades que também nós humanos encontramos quando somos confrontados com as promessas e os perigos associados à Inteligência Artificial: nem mesmo os maiores peritos nestas tecnologias têm conseguido chegar a uma opinião unânime sobre o seu futuro. O conhecimento disponível—os dados utilizados pelo Megatron—podem ser utilizados para argumentar tanto a favor como contra a preservação da Inteligência Artificial. E decidir de que lado está a razão—ou, mais precisamente, em que medida é que cada lado tem razão—torna-se num exercício de futurologia limitado pelo desconhecimento do real potencial da Inteligência Artificial.

"(...) o resultado deste debate é paradigmático das dificuldades que também nós humanos encontramos quando somos confrontados com as promessas e os perigos associados à Inteligência Artificial: nem mesmo os maiores peritos nestas tecnologias têm conseguido chegar a uma opinião unânime sobre o seu futuro."

Na verdade, não é propriamente preciso saber quem tem razão. A incerteza que rodeia a Inteligência Artificial e os desenvolvimentos tecnológicos que o futuro lhe reserva é relativamente inevitável. Ainda que o debate entre quem dá prioridade à maximização dos benefícios da Inteligência Artificial (os "boomers") e quem põe em primeiro lugar a mitigação dos riscos existenciais associados à Inteligência Artificial (os "doomers") possa ser produtivo, dele não se podem esperar respostas definitivas a questões como a que motivou o debate da Oxford Union em 2021. Mas o que nos deve preocupar não é a incerteza que rodeia a Inteligência Artificial—é que essa incerteza possa conduzir à inação. A postura de "esperar para ver" arroga-se de uma neutralidade que não possui: parecendo encerrar um compromisso entre "boomers" e "doomers", é decisivamente enviesada em favor dos primeiros.

D.

Se “esperar para ver” é uma concessão aos “boomers”, e proibir o desenvolvimento de aplicações de Inteligência Artificial é decidir definitivamente em favor dos “doomers”, o que podemos fazer para gerir o impacto da Inteligência Artificial na sociedade moderna? No campo do Direito, a União Europeia decidiu avançar com o primeiro regulamento “horizontal” sobre Inteligência Artificial, propondo-se a regular todas as suas aplicações na medida e em função do risco que criam—mas muito trabalho tem sido também feito a nível setorial, com a regulação de aplicações de Inteligência Artificial em áreas específicas. A importância desse trabalho—reconhecida pelo próprio Regulamento Europeu da Inteligência Artificial—reside nas lições que nos pode ensinar à medida a que a Inteligência Artificial se vai tornando mais prevalente em cada vez mais áreas das nossas vidas.

“No campo do Direito, a União Europeia decidiu avançar com o primeiro regulamento “horizontal” sobre Inteligência Artificial, propondo-se a regular todas as suas aplicações na medida e em função do risco que criam (…).”

Dentro desta lógica, o setor financeiro—e a sua transformação pela Inteligência Artificial—providencia um excelente laboratório: primeiro, para entender melhor os benefícios e riscos criados por produtos e serviços assentes em Inteligência Artificial; segundo, para testar diferentes tipos de respostas regulatórias aos riscos criados pela Inteligência Artificial; e, finalmente, para nos encorajar a repensar o que é que queremos da Inteligência Artificial.

De facto, e a par com meia dúzia de outras áreas—como a saúde e os transportes—o setor financeiro tem-se revelado uma das áreas mais permeáveis à influência da Inteligência Artificial, com países com o Reino Unido a reportar taxas de adesão na ordem dos 70%. Assim se têm então transformado a forma como é prestado aconselhamento financeiro, o modo como se negoceiam instrumentos financeiros em bolsa, e a forma como são feitas as decisões de concessão de crédito. E assim se têm vindo a revelar tanto a importância da Inteligência Artificial para aumentar a eficiência do sistema financeiro, como o perigo que essa tecnologia pode representar para a estabilidade e inclusão financeiras. Do setor financeiro têm saído também algumas das mais criativas soluções regulatórias dirigidas à Inteligência Artificial, desde o enquadramento legislativo europeu aplicável às trocas algorítmicas em bolsa—até hoje o conjunto de regras mais pormenorizado alguma vez aprovado para governar utilização de tecnologias algorítmicas e de Inteligência Artificial—às “caixas de areia regulatórias” em que a indústria financeira tem vindo a testar novas aplicações de Inteligência Artificial em colaboração com reguladores e supervisores.

“Do setor financeiro têm saído também algumas das mais criativas soluções regulatórias dirigidas à Inteligência Artificial (…).”

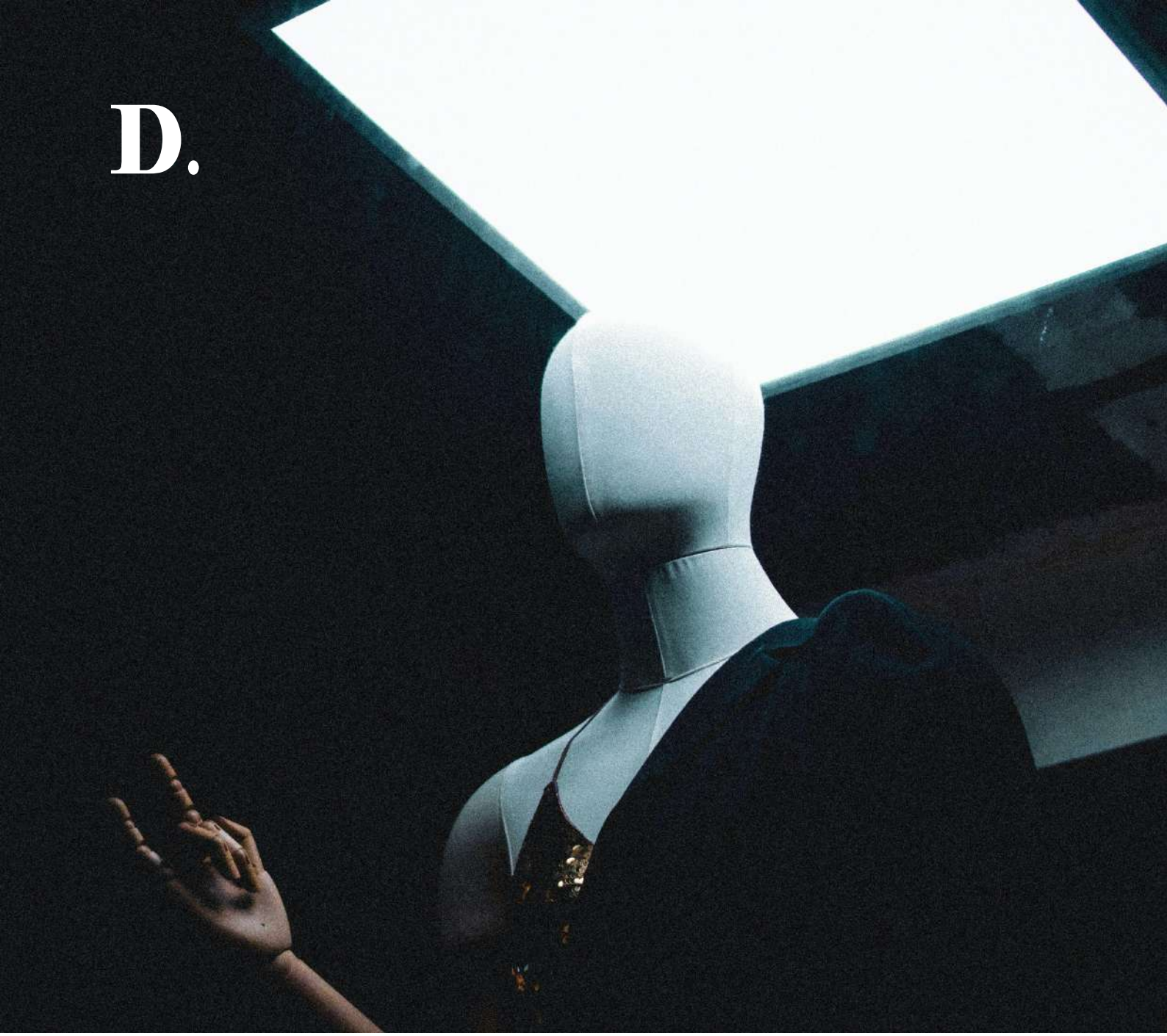
D.

Finalmente, o sistema financeiro convida-nos a debater as expetativas que queremos depositar sobre a Inteligência Artificial. Em discussões sobre a regulação dos sistemas financeiros, a ideia de que tais sistemas devem aspirar não apenas a ser (alocativamente) eficientes, mas também (distributivamente) justos é relativamente consensual: um sistema financeiro que sirva apenas para o enriquecimento desproporcional de poucos à custa de muitos é um sistema financeiro que, em vez de servir a economia, se serve dela. Assim, um olhar sobre o sistema financeiro serve também para nos relembrar de uma pergunta fundamental que urge colocar nesta nova era de Inteligência Artificial: o que queremos nós dela? E é também aqui que o debate sobre a Inteligência Artificial se deve fazer: não interessa apenas discutir se futuros desenvolvimentos tecnológicos nesta área vão trazer os benefícios prometidos pelos "boomers" ou levar à materialização dos riscos antecipados pelos "doomers". Nos debates sobre a Inteligência Artificial—como nos debates sobre o sistema financeiro—importa discutir também como serão distribuídos esses benefícios e riscos, quaisquer que eles venham a ser. Ao contrário dos exercícios de futurologia em que frequentemente redundam os debates sobre Inteligência Artificial, esta é uma discussão que em muito precede as primeiras formas de tecnologia algorítmica. Mas hoje—num contexto de Inteligência Artificial—é mais importante que nunca reavivá-la.

"(...) o sistema financeiro convida-nos a debater as expetativas que queremos depositar sobre a Inteligência Artificial."

Clara Martins Pereira
Assistant Professor, Durham University Law School

D.



INTELIGÊNCIA ARTIFICIAL E MODA
A NOVA ERA DA CRIATIVIDADE

POR MARIA VICTÓRIA ROCHA

D.

Embora haja várias sugestões quanto à delimitação da Inteligência Artificial (IA), é pacífico que abrange o campo das ciências da computação que se centra na criação de programas e mecanismos capazes de adotar comportamentos inteligentes, preparados para aprender a raciocinar como faria um ser humano mediante a imitação das suas funções cognitivas, em que se destacam o pensamento, o raciocínio, a aprendizagem, a tomada de decisões e a resolução de problemas. A IA está a mudar o mundo da criatividade nas indústrias criativas, incluindo na moda. Os designers podem contar com plataformas de IA que de forma quase imediata criam design inovador, projetam padrões exclusivos ou preveem as próximas tendências de moda. Em Nova Iorque, em abril de 2023, ocorreu a primeira "AI Fashion Week", evento onde as centenas de coleções inovadoras apresentadas foram todas geradas por IA.

"A IA está a mudar o mundo da criatividade nas indústrias criativas, incluindo na moda."

A previsão de tendências é a forma mais importante de utilização da IA na moda. As plataformas de IA fazem prospeção de textos e dados para descobrir correlações e padrões em grandes quantidades de dados, fornecendo os elementos de base para a análise preditiva. As redes sociais, websites de comerciantes online e outras páginas web são fontes de dados especialmente relevantes, porque o "scraping" a partir da internet pode revelar como os clientes se sentem em relação a produtos e tendências. Muitas empresas já utilizam este tipo de IA. A previsão tem consequências no trabalho dos designers de moda, orientando-os para uma direção que os consumidores com grande probabilidade acharão apelativa.
A IA também já é estilista virtual. Ao utilizar o histórico de navegação e de compras do consumidor, é possível personalizar e apresentar modelos pré-existentes e disponíveis que também agradarão ao utilizador, bem como sugerir acessórios para uma para uma peça de vestuário em que o utilizador acabou de clicar, atraindo o utilizador para gastar mais dinheiro no site. Todos já experimentos esta funcionalidade ao comprar online. A Alexa, assistente virtual da Amazon, fornece conselhos de moda e ajuda os utilizadores a escolherem o que vestir tendo em conta vários aspetos, como a ocasião, o clima e as tendências de moda.

Muitos designers de moda utilizam a IA como sua assistente. As inovações derivam do acompanhamento próximo das necessidades e desejos do consumidor tendo por base o contexto das suas vidas diárias. A Ivyrevel, agência de moda digital pertencente à H&M, desenvolveu um projeto com a Google denominado *Data_dress*. O objetivo do projeto é proporcionar aos consumidores a possibilidade de terem roupas personalizadas a um custo acessível.

A IA já está a ser utilizada para que se produza a pedido, otimizando a produção e reduzindo o desperdício na indústria da moda. A Amazon tem uma patente para um "sistema de fabrico sob pedido", que utiliza IA para prever tendências de moda e produzir roupas de acordo com a procura,

D.

reduzindo os stocks. A Eon, empresa de tecnologia de moda, usa IA para criar uma "ID Circular" para peças de vestuário. A ID digital ajuda a rastrear o ciclo de vida de uma peça de vestuário, desde a sua produção até ao fim da sua vida útil, incentivando a economia circular e a sustentabilidade.

"A IA já está a ser utilizada para que se produza a pedido, otimizando a produção e reduzindo o desperdício na indústria da moda."

Combinada com a realidade aumentada ou com a realidade virtual, a IA contribui para uma moda digital, largamente desenvolvida durante a pandemia Covid 19. Popularizaram-se novas formas de experimentar e comprar vestuário, calçado e acessórios, podendo as peças ser experimentadas virtualmente antes de serem compradas ou até produzidas fisicamente (cf. os "try on mirrors", os "try on" utilizando o telemóvel, o vestuário e acessórios para avatares, os capacetes ou óculos de realidade virtual com coleções virtuais).

Os "wearables" a as "smart clothes" já são parte do nosso dia a dia.

A IA também pode criar de forma independente com base nos "big data" com que é alimentada e treinada e em resposta aos prompts inseridos pelo utilizador.

As empresas de "fast fashion" e de "ultra fast fashion" vivem da produção e venda de vestuário, calçado e acessórios semelhantes às peças de casas de alta-costura, em muitos casos optam por fazer cópias, de forma muito rápida e a preços muito mais baixos. Por isso, sempre enfrentaram processos de violação de direitos de propriedade intelectual (PI). Beneficiam, todavia, do facto de as legislações aplicáveis poderem ser muito diversas no que toca à proteção pela PI, do facto de se escudarem em tendências, que não são protegidas, e do facto de na indústria da moda haver quem defenda que a cópia é uma forma de estimular a criatividade, pelo que há uma excessiva tolerância. Além disso, a justiça é cara, pelo que não está ao alcance de todos os afetados.

No sistema de PI em que nos inserimos, o objetivo é proteger as criações do ser humano e proporcionar-lhe incentivos para criar. A IA não tem uma "personalidade", nem precisa de incentivos para criar. A criações de IA não podem ser protegidas por direitos de autor ou outros direitos da propriedade industrial, por falta de "personalidade jurídica". Estão, por isso, no domínio público. Todavia, há países, em geral de tradição anglo-saxónica, em que está a ser criado um direito patrimonial, de curta duração, a favor de quem cria as condições para que o "output" surja.

Se tradicionalmente se copiam ou imitam peças de "high fashion" ou de criadores independentes, o jogo está a mudar, as empresas de "fast fashion" e "ultra fast fashion" voltam-se agora contra os seus concorrentes diretos ou contra grandes distribuidores.

D.

Neste novo panorama há um aspeto importante que tem que ver com a IA e que não é exclusivo da moda. Uma ação judicial intentada contra a Shein no ano passado é ilustrativa. Um trio de designers independentes americanos instaurou um processo contra a Shein, em julho de 2023, acusando-a de infringir sistematicamente e em grande escala direitos da PI dos queixosos. Alegadamente, a Shein utiliza um "algoritmo de design" alimentado por IA que identifica e se apropria indevidamente das peças com maior potencial comercial. Embora os pormenores e métodos precisos por detrás do algoritmo da Shein sejam segredos protegidos, é possível inferir certos factos sobre o algoritmo olhando para os seus resultados. O processo da Shein gera frequentemente produtos que são ocasionalmente cópias exatas do trabalho de grandes designers, mas na maioria das vezes, de designers independentes, como os queixosos.

"O processo da Shein gera frequentemente produtos que são ocasionalmente cópias exatas do trabalho de grandes designers, mas na maioria das vezes, de designers independentes, como os queixosos."

Colocando a questão de forma geral, está em aberto saber se quanto uma plataforma de de IA utiliza, sem autorização, obras ou outro material protegido, gerando outputs que são idênticos ou muito semelhantes aos materiais protegidos utilizados, haverá infração de direitos da PI.

É certo que algumas utilizações poderão ser livres, designadamente ao abrigo dos arts. 3º e 4º da Diretiva (UE) 2019/79. A diretiva considera livre a mineração para investigação científica (art.3) e para qualquer finalidade (art. 4). Ao abrigo do art.4, é lícita a mineração para fins comerciais, desde que respeitados os três passos da Convenção de Berna. Todavia, a disposição só se aplica se os titulares de direitos não tiverem reservado expressamente os seus direitos de forma adequada. Noutros países, designadamente nos EUA, podemos questionar quando é que o "fair use" permite a mineração.

É certo que as questões relacionadas com a cópia de direitos de PI na moda existem há muito. No entanto, a IA afeta a escala em que estes problemas se colocam e dá origem a novos complexos problemas.

Maria Victória Rocha

Professora Auxiliar da Faculdade de Direito
(Católica do Porto)

D.

CONSULTURIA 4.0

A REVOLUÇÃO DO SETOR ATRAVÉS DE INTELIGÊNCIA ARTIFICIAL

A transformação digital apresenta-se como o maior catalisador para o dinamismo e competitividade do tecido empresarial moderno, indo além da utilização de tecnologias avançadas. Estende-se ao cerne estratégico das organizações que procuram a inovação, eficiência operacional e dinamismo.

Não fosse a "coragem" e a "excelência" dois dos valores que norteiam a KPMG, numa era onde a tecnologia redefine os próprios limites da excelência, a internalização de ferramentas imbuídas de Inteligência Artificial ("IA") para a otimização das operações torna-se vital para a respetiva concretização.

D.

"(...) numa era onde a tecnologia redefine os próprios limites da excelência, a internalização de ferramentas imbuídas de Inteligência Artificial ("IA") para a otimização das operações torna-se vital para a respetiva concretização."

Neste sentido, várias têm sido as parcerias e desenvolvimentos dinamizados. O compromisso multimilionário entre a KPMG e a Microsoft é o apogeu disso mesmo já que inclui, entre outras, ferramentas como o Azure OpenAI Service e o Copilot.

A título de exemplo, recentemente, o departamento de *Audit Innovation* anunciou recentemente que irá utilizar o Clara AI, ou seja, irá ser utilizada IA no próprio *software* de auditoria.

Adicionalmente, o *Citizen Developer Program* permite a pessoas que não são engenheiros de *software*, desenvolvê-los, o que pressupõe a utilização de ferramentas de automação e IA. Para isso, foi criado um ambiente de desenvolvimento que possibilita os colaboradores de tudo o que é necessário para produzirem as suas *power applications*.

"(...) foi criado um ambiente de desenvolvimento que possibilita os colaboradores de tudo o que é necessário para produzirem as suas *power applications*."

Assim, o papel da área de inovação neste programa é disponibilizar formações e certificações, mantendo um ambiente adequado ao desenvolvimento de aplicações disponível para estimular a criação destas. Isto porque acreditamos veemente que, num futuro mais próximo do que imaginamos, existirá uma grande diferença entre quem sabe utilizar estas ferramentas e quem não sabe.

Desta feita, ao combinar a *expertise* da KPMG com a tecnologia de ponta da Microsoft, têm surgido e/ou sendo adaptadas ferramentas que fomentam a eficiência interna (auxiliando as nossas equipas a prestar um output similar, de uma forma mais rápida), como também a impulsionar a diferenciação da concorrência (pois permitem prestar serviços inovadores através da utilização de tecnologia – com e sem IA). Em particular, a equipa de Tax Innovation está a introduzir a criação de interfaces com os clientes, sendo, por isso, o *output* final um *dashboard* interativo, ao invés do PDF tradicional.

Noutro registo, vários têm sido os eventos internos nesta temática. O mais recente - que decorreu em janeiro e tinha como mote "24 horas de Inteligência Artificial" - permitiu que todos os colaboradores da KPMG a nível global assistissem a formações onde foram abordados desde temas iniciais sobre IA, até como é que a IA poderá ser aplicada em cada um dos departamentos da firma. No fundo, foi um display de tudo aquilo que tem sido desenvolvido nesta área por todas as firmas KPMG.

D.

Já no dia-a-dia, o foco principal subjaz na disponibilização genérica das ferramentas mencionadas, isto é, as ferramentas do *Power Platforms* já são utilizadas quase na sua totalidade. Agora, o objetivo passa por potenciar a utilização genérica de *Power Apps* ou *Power Pages*, o que permitirá uma melhor a implementação de IA, sendo certo que, naturalmente, este avançar deve ser feito com muita cautela, dada a natureza da informação a utilizar. Assim, é expectável que, numa fase inicial, os procedimentos utilizarão dados mais genéricos, disponíveis abertamente na *web* (por exemplo, acórdãos) para, gradualmente, à medida do avançar e eficácia da ferramenta, irem sendo utilizados dados mais sensíveis.

"(...) o objetivo passa por potenciar a utilização genérica de Power Apps ou Power Pages, o que permitirá uma melhor a implementação de IA, sendo certo que, naturalmente, este avançar deve ser feito com muita cautela, dada a natureza da informação a utilizar."

Será que entre 3 a 5 anos o modo de trabalho na firma será diferente? No caso do departamento de Tax, há uma dualidade. Por um lado, existe um desejo de implementação generalizada de IA, principalmente no que concerne a ferramentas que já são utilizadas (daí que o Copilot será, sem margem para dúvidas, uma das primeiras grandes mais-valias). Mas, por outro lado, será necessário testar essas mesmas ferramentas. Isto é, identificar previamente quais é que são as oportunidades, as vantagens e os riscos associados para que, posteriormente, a utilização generalizada seja possível.

"Será que entre 3 a 5 anos o modo de trabalho na firma será diferente?"

Além disso, a própria abordagem a estas ferramentas deve ser efetuada de forma distinta do usual; deve ser um processo de aprendizagem *hands-on*, que é precisamente o que tem acontecido. Por exemplo, para o *Data Snipper* (uma das ferramentas incluída no Excel), oferecemos formação a cada grupo técnico com exemplos práticos específicos destes.

"(...) a própria abordagem a estas ferramentas deve ser efetuada de forma distinta do usual; deve ser um processo de aprendizagem *hands-on* (...)."

Finalmente, o tema do próximo *KPMG Tax Summit* (uma reunião anula onde juntamos centenas dos nossos principais Clientes de diferentes geografias) será, justamente, dedicada aos desafios e oportunidades em matéria de IA. À semelhança do que aconteceu noutros anos, Portugal estará presente neste evento, fazendo-se acompanhar de Clientes que beneficiarão de diversas intervenções de especialistas e da a apresentação de ferramentas *made in KPMG Portugal*.

Da teoria às pessoas:

Mas qual é, no final de contas, o leque de ferramentas que estamos a utilizar? Existem diversos serviços incluídos no Azure AI da Microsoft, para a criação de aplicações com funcionalidades de IA.

Em termos genéricos, podemos classifica-los como *AI Tools* (onde se incluem o Azure AI Studio, o Copilot e o Azure AI Search), *Computer Vision Models* (onde se inclui o Azure Document Intelligence), Learning Models (Azure Machine Learning), *Language Models* (Azure AI Language e Azure OpenAI Service – GPT) e outros serviços de AI não categorizados (onde se incluem o Azure AI Speech, Azure AI Translator, Azure AI Content Safety, Azure AI Vision e Azure OpenAI Service – DALL-E).

Destes, aqueles que diria que se têm mostrado mais estimulantes, para mim, são as ferramentas de *prompt engenering*, cada vez mais empregues na diligência de clientes. Isto é, estamos a introduzir IA para implementar modelos que analisam comportamentos, para além de analisar simplesmente as transferências monetárias, com o intuito de avaliar se as transações são ou não consistentes, dado que tais situações são passíveis de gerar falsos positivos.

"(...) estamos a introduzir IA para implementar modelos que analisam comportamentos, para além de analisar simplesmente as transferências monetárias, com o intuito de avaliar se as transações são ou não consistentes, dado que tais situações são passíveis de gerar falsos positivos."

Pedro Penedo, Ana Rodão e José Afonso

Partner de Innovation, Director de Tax Innovation
e Senior Advisor Technology Consulting - KPMG

D.

ENTRE DOOMERS E UTOPIANS

O FUTURO DA IA E DA HUMANIDADE

Quando se fala em Inteligência Artificial (IA) e nas suas potencialidades transformadoras da sociedade, o primeiro exemplo habitualmente citado é o ChatGPT. Este faz parte de uma onda de novas ferramentas de Inteligência Artificial Generativa (IAG), que se definem pela sua capacidade de gerar conteúdo novo (p. ex., texto, imagem, vídeo, áudio) a partir de um input ou prompt.

A IAG tem feito manchetes e sido notícia de abertura de telejornais pela sua mimetização capaz e surpreendente de características associadas ao ser humano, tais como o domínio da linguagem, a criatividade ou a inteligência. Ainda que se possa dizer que a IAG, e em particular o LLM, nada mais é que um papagaio estocástico (Bender et al., 2021), porque, não possuindo compreensão do significado das palavras que transmite, se limita a reproduzir padrões observados no corpus que lhe serviu de treino, não são despiciendos os riscos associados à sua utilização.

"Ainda que se possa dizer que a IAG (...) nada mais é que um papagaio estocástico (...), porque, não possuindo compreensão do significado das palavras que transmite, se limita a reproduzir padrões observados no corpus que lhe serviu de treino, não são despiciendos os riscos associados à sua utilização."

Riscos que, para Yuval Noah Harari (2023), advêm do facto de a IA ter "hacked the operating system of our civilisation". Para o Autor, quando um sistema de IA é melhor que um "humano médio a contar histórias, compor melodias, desenhar imagens e escrever leis e escrituras" tornam-se imagináveis cenários em que campanhas eleitorais são inundadas de fake news, novos cultos são formados com base em escrituras geradas por IA ou a opinião pública é manipulada em tópicos sensíveis como o aborto. Curiosamente, e um tanto ou quanto paradoxalmente, o risco de concretização destes cenários torna-se mais credível quanto mais fácil se tornar a utilização massiva de sistemas de IA para o estabelecimento de relações comunicativas individuais entre a IA e um ser humano específico, marcadas pela intimidade, empatia e atenção. Artificiais, sintéticas, sim. Antropomorfo-logicamente reais, também. Trata-se do risco da falsa intimidade para que nos alerta Yuval Noah Harari. Um risco que, para o historiador e filósofo, pode levar ao fim da história humana.

D.

"(...) o risco de concretização destes cenários torna-se mais credível quanto mais fácil se tornar a utilização massiva de sistemas de IA para o estabelecimento de relações comunicativas individuais entre a IA e um ser humano específico, marcadas pela intimidade, empatia e atenção."

Porém, um outro tipo de risco terá motivado centenas de cientistas, académicos, políticos e executivos a assinar, em maio de 2023, uma curta, mas impactante declaração: "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war". Na sua base está supostamente o receio do surgimento de uma Inteligência Artificial Geral (IAG), também conhecida por Inteligência Artificial Forte ou Plena: um sistema de IA com um nível de inteligência idêntico ou superior ao do ser humano, com autonomia e capacidade de auto-aprendizagem. Um sistema de IA que poderia mesmo tornar-se senciente e determinar o fim da humanidade. Seja por maldade, seja por competência - os seres humanos serão meros obstáculos à concretização de objetivos a atingir pela AGI e a sua eliminação aparece como um passo lógico. A última destas alternativas foi sugerida por Stephen Hawking em 2015, num AMA no Reddit: "You're probably not an evil ant-hater who steps on ants out of malice, but if you're in charge of a hydroelectric green energy project and there's an anthill in the region to be flooded, too bad for the ants. Let's not place humanity in the position of those ants".

"Seja por maldade, seja por competência - os seres humanos serão meros obstáculos à concretização de objetivos a atingir pela AGI e a sua eliminação aparece como um passo lógico."

Talvez os cenários apocalíticos, à guisa do Exterminador Implacável, sejam fruto de uma visão demasiado pessimista do futuro da humanidade. Do nosso futuro. Ou talvez todos os problemas da humanidade possam ser solucionados com a IAG e com os inevitáveis avanços científicos, económicos e sociais por si proporcionados - uma ideologia batizada com o nome A.G.I.smo. por Evgeny Morozov, e por si criticada por andar de mãos dadas com o neoliberalismo. Talvez a virtude tenha, como muitas vezes, de ser descoberta na moderação, no caminho intermédio, entre os pessimistas (doomers) e os utopistas (utopians). Mas sem dúvida que os limites do avanço tecnológico têm de ser definidos não apenas por engenheiros, mas também por juristas...

Pedro Miguel Freitas

Professor Auxiliar da Faculdade de Direito
(Católica do Porto)

D.

INTELIGÊNCIA ARTIFICIAL, UM DILEMA LIBERAL

João Cotrim de Figueiredo

D.

Se pensarmos bem, esta mais recente vaga de debate sobre a inteligência artificial devia suscitar alguma surpresa. Afinal, a tecnologia e grande parte das suas aplicações mais úteis já existem há uma década, ou mais. Porque é que se intensificou então o debate?

Em termos cronológicos, a razão remonta a 30 Novembro de 2022, data do lançamento do Chat GPT, um chatbot desenvolvido pela empresa OpenAI. A sua capacidade e velocidade de discurso e 'raciocínio' pronto-humano apanharam o Mundo de surpresa. Pela primeira vez, uma máquina poderia passar o famoso Teste de Turing que o cientista britânico idealizou em 1950.

Num artigo científico, Alan Turing descrevia um conjunto de regras a que chamou The Imitation Game (sim, o que deu nome ao filme) que definia as condições em que se poderia dizer que uma máquina possuía inteligência humana. Basicamente, as condições eram que se, através de um diálogo escrito, mantido à distância entre uma máquina e uma pessoa, esta última não tivesse mais do que 30% de probabilidade de identificar o interlocutor como não-humano, poderia dizer-se que a máquina possuía inteligência de tipo humano.

Com o Chat GPT, e outros chatbots que apareceram nos meses seguintes, ficou, pela primeira vez, evidente que as máquinas conseguiam exibir raciocínio humano. Com a incrível velocidade de evolução dos modelos de linguagem subjacentes aos GPTs, ficou, pela primeira vez, patente que as máquinas poderiam superar a capacidade humana num horizonte não muito distante. Foi isso, que desencadeou a recente torrente de entusiasmo, e outra de catastrofismo, à volta da inteligência artificial.

"Com o Chat GPT, e outros chatbots que apareceram nos meses seguintes, ficou, pela primeira vez, evidente que as máquinas conseguiam exibir raciocínio humano".

A torrente de entusiasmo é fácil de entender. É uma enorme façanha do engenho e da criatividade humanas desenvolver uma tecnologia que nos poderá facilitar a vida em tantos e tantos domínios da nossa vida. Da saúde à educação, da agricultura à indústria, do software à cibersegurança do setor financeiro aos serviços jurídicos, é difícil encontrar uma área de atividade humana que não possa vir a ser disruptivamente alterada pela inteligência artificial.

Para quem acha que o Siri, a Alexa, as recomendações automáticas, o preenchimento de texto, as viaturas de condução autónoma, o reconhecimento facial, a ativação por voz em linguagem natural são funções quase milagrosas da nossa atual tecnologia, deve preparar-se para evoluções ainda mais incríveis e imprevisíveis no futuro próximo.

D.

A torrente de catastrofismo é, igualmente, fácil de entender. A velocidade exponencial - e, em última análise - descontrolada do desenvolvimento das capacidades cognitivas das máquinas, corresponde, com exatidão, ao conceito de Singularidade que alimentou boa parte da ficção científica da segunda metade do século XX. A Singularidade, em termos simples, é aquele momento da História em que a inteligência artificial superará a inteligência humana, com consequências imprevisíveis quanto ao significado da condição do ser humano e, mesmo, quanto à sua sobrevivência.

"A velocidade exponencial (...) descontrolada do desenvolvimento das capacidades cognitivas das máquinas, corresponde, com exatidão, ao conceito de Singularidade que alimentou boa parte da ficção científica da segunda metade do século XX".

O cientista americano Ray Kurzweil popularizou o conceito de Singularidade no seu livro "The Singularity is Near" de 2005, em que estimou que a data desse evento ominoso seria em 2045. Dadas as evoluções mais tardias, até pode ser que Kurzweil tenha pecado por excesso. Talvez lá cheguemos mais cedo.

Os catastrofistas não são só os tradicionais "velhos do Restelo" ou arautos da desgraça. Figuras ponderadas e respeitadas como o falecido Stephen Hawking, Bill Gates, Elon Musk ou Eric Weinstein já manifestaram publicamente os seus receios quanto às consequências do desenvolvimento não-controlado desta poderosa tecnologia.

Muitas destas personalidades advogam maior ponderação no desenvolvimento das aplicações de inteligência artificial, não apenas pelos riscos mais extremos como a extinção da espécie humana, mas também por consequências de curto prazo ao nível do desemprego de pessoas qualificadas, ou a facilidade de produção de desinformação (através de deepfakes, por exemplo), ou os riscos para a privacidade dos dados pessoais, ou ao nível dos viés ("biases") que os modelos de linguagem ("large language models") que são os vastos conjunto de dados nos quais os GPTs ("generative pre-trained") são treinados possam carrear para os bots finais.

Para um liberal como eu, esta matéria coloca um difícil dilema.

Por um lado, o desenvolvimento da inteligência artificial é um monumento à criatividade e à inteligência humanas e tal é, também, a confirmação da tese liberal de que os indivíduos livres são capazes de responder a estímulos e a necessidades, e que tal resposta conduz à nossa evolução enquanto espécie, mas também enquanto sociedade. É bom termos a noção que o desenvolvimento da inteligência artificial necessitou de grandes avanços e inspirações individuais, bem como de enorme colaboração entre especialistas de áreas muito distintas.

D.

Áreas técnicas e tecnológicas como estatística, pesquisa operacional, otimização matemática, redes neurais artificiais, arquitetura de processadores GPUs (Graphics Processing Unit) assistiram a tremendos desenvolvimentos. Só para terem uma ideia do impacto económico da evolução da inteligência artificial, só nesta tecnologia de GPUs, o mercado é liderado pela NVidia, uma empresa americana que em janeiro de 2024 vale cerca de $1,4 bn, um crescimento de 150x nos últimos 10 anos!

No entanto, estas áreas mais técnicas tiveram de recorrer a conhecimentos de áreas tradicionalmente conotadas com as chamadas Humanidades. Sem valiosos insights de disciplinas como Psicologia, Linguística, Filosofia e Neurociência não teriam sido possíveis os enormes avanços na inteligência artificial. Esta capacidade dos seres humanos juntarem as suas capacidades individuais em diversos domínios e colaborarem em prol do bem comum é algo que um liberal acredita ser um dos grandes motores do bem-estar e do desenvolvimento humanos.

Mas, por outro lado, e é um lado que não devemos levianamente ignorar, é possível que esta evolução tecnológica não seja igual a tantas outras anteriores. Pode mesmo ser que estejamos numa via sem precedentes, em que o princípio da liberdade individual, que os liberais consideram o valor político mais importante, possa estar em perigo. Como?

"Pode mesmo ser que estejamos numa via sem precedentes, em que o princípio da liberdade individual, que os liberais consideram o valor político mais importante, possa estar em perigo".

Se as nossas ações forem determinadas por informação falsa ou errónea produzida por máquinas inteligentes, não seremos livres. Se as nossas intenções forem contrariadas ou impedidas por máquinas que tenham intenções próprias, não seremos livres. Digo mais: se, enquanto indivíduos, não tivermos a certeza de que tais cenários não são possíveis, não seremos livres.

Lamento não ter respostas para muitas destas questões. Mas não lamento ter uma série de perguntas incómodas para as quais é fundamental termos respostas. Todos não seremos demais. A começar por mim. E por vocês que me lêem.

João Cotrim de Figueiredo

Deputado à AR pela Iniciativa Liberal

D.

INTELIGÊNCIA ARTIFICIAL E POLÍTICA

POR WILLIAM HASSELBERGER E INÊS GREGÓRIO

Muitos analistas preveem que a Inteligência Artificial esteja próxima de se tornar uma tecnologia universal semelhante à eletricidade: um tipo de tecnologia com a capacidade de influenciar, ou mesmo de transformar fundamentalmente, todas as esferas da vida humana. A vida familiar, o trabalho, as artes, a ciência, a medicina, a educação, a política e o governo estão a ser progressivamente remodelados pela Inteligência Artificial e tecnologias relacionadas, como ferramentas de vigilância digital e robótica. Alguns investigadores e futuristas de Inteligência Artificial vão mais longe e afirmam mesmo que estamos a caminhar em direção à Inteligência Geral Artificial de "nível humano", ou sistemas de Inteligência Artificial com capacidades cognitivas que igualam ou excedem os humanos em todos os domínios.

"A vida familiar, o trabalho, as artes, a ciência, a medicina, a educação, a política e o governo estão a ser progressivamente remodelados pela Inteligência Artificial e tecnologias relacionadas (...)."

Os governos, as instituições académicas e as empresas tecnológicas estão cada vez mais conscientes da necessidade de desenvolver quadros normativos para determinar como deve a Inteligência Artificial ser aplicada na sociedade de uma forma que seja eticamente responsável, confiável, benéfica e humana. Na política internacional, a Presidente da Comissão Europeia, Ursula von der Leyen, utiliza a linguagem da Universidade de Stanford e dos laboratórios do MIT para

D.

argumentar que a Inteligência Artificial deve ser "centrada no ser humano", enquanto o Parlamento Europeu aprovou já o texto inicial da Lei da Inteligência Artificial, que categoriza diferentes utilizações da Inteligência Artificial em termos dos seus supostos riscos para os seres humanos: Risco Inaceitável, Risco Alto e Baixo Risco, cada categoria implicando diferentes níveis de regulação e supervisão.

A questão de como utilizar adequadamente a Inteligência Artificial na sociedade está para lá do domínio das questões técnicas e transporta-nos para a área controversa do pensamento normativo, ou da ética, que se debruça sobre o que é bom e mau, o que é justo e injusto, sobre o que significa agir em relação aos outros de forma responsável ou irresponsável. Por isso, o desenvolvimento tecnológico de ponta, para ser um benefício e não um dano, não pode evitar os antigos debates da filosofia, da política e do direito sobre como devemos agir, viver e ordenar as nossas sociedades.

"(...)o desenvolvimento tecnológico de ponta, para ser um benefício e não um dano, não pode evitar os antigos debates da filosofia, da política e do direito (...)."

Uma das aplicações práticas da Inteligência Artificial é a sua utilização para automatizar a tomada de decisões e o desempenho de tarefas cognitivas associadas à inteligência humana. Um exemplo bem conhecido é a Inteligência Artificial generativa, baseada em "large language models" (LLMs), que tem a capacidade de produzir conteúdo verbal e visual semelhante ao humano, muitas vezes de qualidade muito elevada. As potencialidades de uso fraudulento da Inteligência Artificial generativa são várias e incluem a sua utilização por parte de atores hostis ao estado para produzirem desinformação em massa e propaganda de alta qualidade, hiper-realista e personalizada. Particularmente preocupante é o impacto destas tecnologias nas próximas eleições em 2024 em países como os EUA, a Grã-Bretanha, a Índia, o México, a Indonésia e Taiwan. As agências de informação e segurança nacionais terão de ser altamente competentes na identificação, rastreio e, em alguns casos, no encerramento dos "exércitos de robôs" que estão a ser treinados para interferir no processo político nos países democráticos. Aqui, será necessário utilizar "bons" sistemas de Inteligência Artificial para identificar as atividades de "maus" atores de Inteligência Artificial, como bots financeiros fraudulentos, redes terroristas que utilizam Inteligência Artificial para desenvolver planos e armas, e redes de propaganda que distorcem as esferas mediáticas e os processos políticos.

"As potencialidades de uso fraudulento da Inteligência Artificial generativa são várias e incluem a sua utilização por parte de atores hostis ao estado para produzirem desinformação em massa e propaganda de alta qualidade, hiper-realista e personalizada."

D.

Uma outra utilização da Inteligência Artificial diz respeito à análise e previsão de dados, onde o sistema toma como input uma imagem e gera, como output, algum tipo de classificação ou previsão. Vários países têm utilizado a visão computacional de Inteligência Artificial para funções de vigilância biométrica generalizada da parte das autoridades públicas, como reconhecimento facial e rastreamento de matrículas de veículos. No entanto, A Lei da Inteligência Artificial que está a ser elaborada pela Parlamento Europeu, na sua forma atual, proíbe a maioria das utilizações de Inteligência Artificial pelas autoridades públicas para identificação biométrica em tempo real, por se enquadrar na categoria de "Risco Inaceitável" (exceto em circunstâncias especiais com aprovação judicial). Até porque a visão computacional e a digitalização facial por Inteligência Artificial podem ir muito mais além da vigilância e identificação de pessoas e ser utilizados para, com base apenas na imagem do seu rosto, fazer previsões sobre o estado emocional e mental de uma pessoa, a sua situação clínica, o seu carácter moral, a sua ética de trabalho ou a probabilidade de criminalidade futura, entre outras características. Aqui, os dilemas éticos tornam-se ainda mais evidentes.

Os recentes avanços em sistemas de Inteligência Artificial apanharam de surpresa mesmo os seus criadores. As sociedades democráticas têm, por isso, a responsabilidade, mas também um período de tempo limitado, para garantir que a aplicação da Inteligência Artificial na sociedade seja ética, benéfica e humana. A análise utilitarista de custo-benefício não é suficiente. E o "AI Act" da UE, embora seja um passo histórico, é talvez demasiado amplo na sua análise, demasiado restritivo em alguns pontos e demasiado focado na tecnologia em si, em vez de se concentrar no impacto das diferentes aplicações da Inteligência Artificial nos utilizadores humanos e nas comunidades em áreas específicas da vida. As instituições públicas terão de colaborar com a indústria, com investigadores académicos e com especialistas em ética para articular quadros normativos adequados que possam avaliar e orientar o desenvolvimento e a aplicação da Inteligência Artificial na sociedade.

"Os recentes avanços em sistemas de Inteligência Artificial apanharam de surpresa mesmo os seus criadores."

William Hasselberger e Inês Gregório

Professor Associado do IEP
Professora Auxiliar Convidada do IEP

D.

A INTELIGÊNCIA ARTIFICIAL NO CONTEXTO DA UNIÃO EUROPEIA

PELA COMISSÃO EUROPEIA, REPRESENTAÇÃO EM PORTUGAL

A Inteligência Artificial (IA) traz promessas e desafios às nossas sociedades e economias. A difusão de redes digitais de alta velocidade, a disponibilidade de grandes volumes de dados e a capacidade de os processar com supercomputadores contribuíram para o desenvolvimento de software "inteligente", utilizado em vários sectores económicos – da saúde à mobilidade urbana, dos média às finanças, etc.

Os fundamentos científicos e tecnológicos da inteligência artificial, que estão na base das recentes inovações disruptivas, aconteceram há cerca de 80 anos. Os trabalhos de cientistas como Alan Turing, com o seu famoso artigo sobre um hipotético teste de IA publicado em 1950, conhecido por "*imitation game*", ou de Claude Shannon que em 1951 construiu o rato "Theseus", um pequeno robot que aprendia a orientar-se num labirinto, são exemplos de importantes passos no conhecimento teórico e aplicado em IA após a segunda guerra mundial. Muitos outros passos se seguiram, incluindo a investigação em redes neuronais inspiradas no sistema nervoso de mamíferos, insetos, etc., abrindo novas áreas do saber.

A crescente maturidade da IA permite ir além da automação industrial baseada em robótica, que acarreta por si só fortes impactos socio-económicos. São exemplo os algoritmos sofisticados gerando "diálogos fluentes humano-máquina" (IA generativa) oferecidos como serviço através da Internet, que resultam de décadas de investigação em ciência da computação e de dados, e de engenharia de ponta. A divulgação do acesso, sem restrições, a essas aplicações através da web aumentou a sensibilização pública e desencadeou vivas discussões sobre o seu enquadramento e impactos.

"A crescente maturidade da IA permite ir além da automação industrial baseada em robótica, que acarreta por si só fortes impactos socioeconómicos".

D.

Há dois aspetos a salientar na estratégia da Comissão Europeia para a IA. Por um lado, a proteção dos cidadãos e das nossas sociedades requer uma constante reflexão política e o desenho de regimes jurídicos adequados aos novos tempos. Por outro, sabendo que os avanços em IA têm uma dimensão global, é importante assegurar condições para a competitividade das nossas economias liderando o desenvolvimento científico e tecnológico.

Em relação ao primeiro aspeto, a UE quis dar um forte sinal político com o acordo sobre o Regulamento IA de dezembro de 2023, entre o Parlamento Europeu e o Conselho, e com a revisão de um plano coordenado para a IA. A presidente da Comissão, Ursula von der Leyen, destacou a natureza transformadora da IA na vida dos cidadãos e sublinhou a importância deste acordo que é o primeiro quadro jurídico abrangente a nível mundial em matéria de IA, salvaguardando os nossos valores europeus.

A segurança e os direitos fundamentais dos cidadãos em áreas como a privacidade, a não discriminação, ou a atendibilidade da informação faz com que a exigência aumente em relação aos novos sistemas inteligentes que devem cumprir normas rigorosas em linha com as aplicáveis às tecnologias tradicionais, previstas no direito europeu. Ao centrar a regulamentação nos riscos identificáveis "a priori", o acordo promoverá a inovação responsável ambicionando garantir a segurança das pessoas e das empresas e apoiará o desenvolvimento e a adoção de uma IA fiável na UE.

"A segurança e os direitos fundamentais dos cidadãos (...) faz com que a exigência aumente em relação aos novos sistemas inteligentes (...)".

A Comissão está a preparar a aplicação do novo Regulamento, incluindo a criação de capacidades institucionais para o desenvolvimento e utilização de IA segura e ética. Irá apoiar as administrações públicas da UE através de um novo serviço para supervisionar e coordenar a aplicação de sistemas de IA, alinhando as políticas ao nível da Comissão, organismos da UE, Estados-Membros e partes interessadas.

Em relação aos desafios para a competitividade das nossas economias, não basta ter estado a um bom nível nos últimos 30 anos em infraestruturas avançadas de conectividade e computação. A base de conhecimento científico e tecnológico é de elevada qualidade em muitos países da EU, mas há riscos de que os novos líderes globais em IA sejam predominantemente de outros continentes, tendo em conta o cenário atual de domínio nas redes digitais por parte das grandes empresas da "*web economy*". A UE deve reforçar a sua capacidade de inovação em IA e software avançado, e está a fazê-lo com investimentos dos programas Europa Digital e Horizonte Europa, em suporte a centros de excelência, universidades e através de uma parceria de I&D com indústria inovadora.

D.

A UE quer afirmar-se na vanguarda seguindo a ambição expressa pela presidente Ursula von der Leyen no seu último discurso do estado da União de setembro de 2023, de promover ações concretas que facilitem o acesso à rede europeia de supercomputadores por parte de "*startups*" e PMEs inovadoras. O ano de 2024 começou com o anúncio da vice-presidente executiva da Comissão responsável pela transição digital, Margrethe Vestager, e do Comissário para as áreas da industria e espaço, Thierry Breton, lançando um novo pacote em suporte para IA avançada. É um passo importante para dinamizar um tecido com enorme potencial de inovação, que pode estabelecer precedentes mundiais para uma tecnologia IA de elevados padrões de ética. Neste contexto, a parceria EuroHPC para a computação de alto desempenho financiada pelos programas Horizonte Europa e Europa Digital lançará um novo pilar de «fábricas de IA» em torno das instalações de supercomputação, num valor que pode atingir 600 milhões de Euros até 2027.

"A UE quer afirmar-se na vanguarda seguindo a ambição expressa pela presidente Ursula von der Leyen (...)".

Recorde-se que Portugal – através do "Minho Advanced Computing Center (MACC)" – tem um dos 8 supercomputadores da rede EuroHPC, financiado pela UE e pela Fundação para a Ciência e Tecnologia, colocando a Europa na liderança à escala global, estando previstos 2 novos supercomputadores com capacidade à escala "*exa*" (1 milhão de milhão de milhão de operações por segundo, exa-FLOP) em 2024- 25. Projetos para estabelecer plataformas IA para a reutilização de medicamentos reduzindo o tempo e os custos de desenvolvimento, ou para criar um Centro de Excelência de análise de células e sistemas in vitro combinados com abordagens computacionais, são exemplos de colaborações em AI de ponta, com financiamento Europeu envolvendo parceiros portugueses e de diferentes Estados-Membros.

Portugal apresenta uma boa participação de entidades de I&I em projetos de IA financiados pelo Horizonte Europa. Registou-se entre 2021 e 2023 o envolvimento de 144 entidades em 181 projetos, dos quais 35 coordenados por um parceiro português. O financiamento atinge os 112 M€, tendo já ultrapassado o total dos 7 anos do H2020 (202 projetos, 108 entidades e 92 M€). O sector empresarial é responsável pela captação de quase 30% do orçamento. Portugal faz também parte da rede europeia de polos de inovação digital com 17 "*digital innovation hubs*" (DIHs) sendo um deles focado em IA e ciência dos dados (ATTRACT DIH) e 15 identificaram a IA com uma das competências tecnológicas principais.

A corrida internacional para a liderança em IA está a intensificar-se. A UE não pode ser um ator passivo e correr o risco de ser marginalizada nas futuras descobertas cientificas e tecnológicas. Como referido pela Comissária Iliana Ivanova com a pasta da Inovação, Investigação, Cultura, Educação e Juventude, assegurar uma investigação científica de ponta é essencial para a prospe-

D.

ridade e soberania tecnológica da Europa, para garantir a sua segurança e para defender os seus valores.

"A corrida internacional para a liderança em IA está a intensificar-se. A UE não pode ser um ator passivo e correr o risco de ser marginalizada nas futuras descobertas cientificas e tecnológicas".

Os desafios da IA para a "fábrica da ciência europeia", como a preservação da integridade científica, prevenção da utilização indevida da tecnologia, utilização de supercomputadores e dos muitos dados disponíveis fazem parte do parecer do Mecanismo de Aconselhamento Científico Europeu (SAM) a ser publicado nos próximos meses, baseado no artigo de prospetiva de Julho de 2023. A União Europeia sairá fortalecida promovendo um debate transparente e participado que acompanhe investimentos em I&D&I sobre Inteligência Artificial, respeitadoras dos valores Europeus e que suportem a transição ecológica e digital das nossas sociedades e economias.

Carlos Morais Pires, António Vicente e Sofia Moreira de Sousa

Comissão Europeia
(Representação em Portugal)

D.

APLICAÇÃO DA IA À SAÚDE

BREVE APONTAMENTO TÉCNICO

PELO CONSELHO NACIONAL DE ÉTICA PARA AS CIÊNCIAS DA VIDA

O Conselho Nacional de Ética para as Ciência da Vida (CNECV) vem reflectindo, no âmbito das suas competências, sobre os potenciais impactos da utilização da Inteligência Artificial (IA) no domínio das ciências da saúde. Neste contexto, procedeu a uma sistematização de cinco planos estruturantes de intervenção da IA, tendo analisado as implicações sociais mais impactantes e elaborado recomendações no sentido de potencializar os benefícios e mitigar os prejuízos.

"O **Conselho Nacional de Ética para as Ciência da Vida (CNECV) vem reflectindo, no âmbito das suas competências, sobre os potenciais impactos da utilização da Inteligência Artificial (IA) no domínio das ciências da saúde.**"

O primeiro plano é o da investigação biomédica em que se atende ao impacto do reconhecimento de padrões em quantidades gigantescas de dados (Big Data), a partir do que um volume crescente de informação ganha sentido e gera conhecimento, isto é, adquire valor. Exigindo-se sempre a privacidade e a confidencialidade dos dados, destacaram-se duas questões éticas: a da revisão dos critérios de propriedade intelectual, estabelecendo requisitos claros e transparentes para uma determinação rigorosa de autoria e de contribuição, em todas as fases da investigação, a par ainda da revisão da atual lei das patentes que o progresso da IA torna obsoleta; e a da necessidade de, perante a resposta unívoca da IA, manter a formulação de dúvidas e hipóteses científicas, de implementar diferentes metodologias explicáveis e de proceder à replicação dos estudos, promover a originalidade das análises e das interpretações das realidades em estudo.

D.

O segundo é o da assistência clínica, em que se privilegia o impacto da assistência digital – através de chatbots, websites, apps – muito útil no atendimento remoto e triagem de pacientes, na elaboração de diagnósticos e na recomendação de terapêuticas, libertando os recursos humanos para tarefas mais complexas. Contribui para uma maior qualidade e acessibilidade a cuidados de saúde, introduzindo, porém, um afastamento entre as pessoas. Assim, do ponto de vista ético importa, promovendo a digitalização dos serviços de saúde em rede integrada e extensiva ao território nacional, atender ao potencial impacto da mediação tecnológica no distanciamento das relações pessoa-a-pessoa (a relação dialogante entre os pacientes, utentes e famílias e os profissionais de saúde constrói a confiança indispensável no processo terapêutico), não deixando de disponibilizar vias alternativas de comunicação. Realça-se ainda que, através das aplicações validadas de IA, se reforça tanto a autonomia como a responsabilidade da pessoa na gestão da sua saúde.

"(...) do ponto de vista ético importa, promovendo a digitalização dos serviços de saúde em rede integrada e extensiva ao território nacional, atender ao potencial impacto da mediação tecnológica no distanciamento das relações pessoa-a-pessoa".

O terceiro plano é o da gestão hospitalar e os benefícios da intervenção clínica à distância, como a telessaúde. Esta é um factor de potenciação de equidade no acesso à saúde, na medida em que permite uma gestão eficaz e uma prestação contínua de cuidados (sem limitações geográficas e inclusiva de cidadãos com menor mobilidade), promovendo a sustentabilidade. Deve, pois, ser impulsionada, a par da necessária agilização do fluxo de informação para benefício dos doentes, a qualificação dos cuidados e a otimização dos serviços, protegendo os acessos, através de uma rigorosa codificação de dados, não deixando de garantir que o factor humano se mantém presente na programação de sistemas de IA e em todos os processos de tomada de decisão.

"O terceiro plano é o da gestão hospitalar e os benefícios da intervenção clínica à distância, como a telessaúde. Esta é um factor de potenciação de equidade no acesso à saúde (...)".

A administração da saúde pública constitui o quarto domínio sistematizado, em que se destaca a importância da codificação de dados, a partir de sistemas de classificação internacionalmente reconhecidos, para uma interpretação fiável e consequente em termos, por exemplo, de vigilância epidemiológica e deteção precoce de surtos. Importa, pois, eticamente, investir num padrão harmonizado e global de codificação, num quadro de governança comum, assim rentabilizando os dados pessoais de saúde em prol da saúde pública, mas sempre no respeito pelos direitos humanos. Deve-se também contrariar vieses conducentes à discriminação de grupos populacionais e os riscos da colonização de dados potencializadores do aprofundamento das assimetrias de poder que estabelece.

D.

O último plano estruturado é o do ensino e educação em saúde e o impacto da realidade virtual, sob a dupla perspectiva do profissional e do utente. Valorizam-se as metodologias de ensino e de formação profissional contínua assistidas pela IA, desde que validadas em evidência científica, e exorta-se ao acesso e utilização competente dos sistemas de IA para promover a educação em saúde, numa atitude informada e crítica por parte do cidadão, compreendendo o funcionamento da IA e consciencializando os riscos associados.

Brevemente, a IA só será benéfica se estiver ao serviço das finalidades humanas, devendo, para tal: alicerçar-se nos princípios éticos identitários do humano, no respeito pelos seus valores nucleares e direitos fundamentais; e centrar-se sempre no humano (uma lógica human-centered design em oposição à user-centered design), isto é, rejeitando o pragmatismo da captura por interesses sectários e/ou particulares, mantendo-a como instrumento de realização do humano, individual e socialmente considerado.

"(...) a IA só será benéfica se estiver ao serviço das finalidades humanas, devendo, para tal: alicerçar-se nos princípios éticos identitários do humano, no respeito pelos seus valores nucleares e direitos fundamentais; e centrar-se sempre no humano (...)."

Maria do Céu Patrão Neves, Miguel Ricou e Inês Godinho

Presidente e Membros do CNECV

D.

D.

INTELIGÊNCIA ARTIFICIAL E ÉTICA

De forma a que se possa tratar, mesmo de modo sumário, o tema «inteligência artificial (IA) e ética», há que definir os termos fundamentais: «inteligência» e «ética». Por «inteligência», entendemos o acto lógico de surgimento do sentido, acto que, sem outra contribuição, constitui o que é o cerne do ser humano como *entidade*, precisamente, *de sentido*: sem este acto, nada surge em termos de sentido, não podendo, assim, haver algo como «ser humano»; sem este acto, nada é referenciável.

Por «ética», entendemos, não qualquer forma adjectiva, como, por exemplo, surge em afirmações como 'uma política ética', mas, substantivamente, todo o âmbito próprio da interioridade lógica-espiritual humana, que se constitui através do exercício da capacidade de escolha – em seu absoluto, que sempre supera a relatividade contextual –, cuja actualização permite ao ente humano ser em acto próprio e irredutível. Ética é, deste modo, o acto da interioridade electiva que constitui o ser humano como realidade autónoma e irredutível, salvo aniquilação, física ou espiritual. «Ética» não se confunde com «deontologia», com «direito», com qualquer forma de normatividade, sempre de índole nomológica, que transcende a pura interioridade espiritual humana, único foro ético.

"Ética é (...) o acto da interioridade electiva que constitui o ser humano como realidade autónoma e irredutível, salvo aniquilação, física ou espiritual".

A questão da inteligência é uma questão ontológica, diz respeito ao *ser do acto de intuir*. Ora, não podendo haver ser humano sem este acto, compreende-se (intui-se) que é a inteligência em acto que cria o ente humano como propriamente humano, como ser lógico, de «logos», de sentido propriamente espiritual.

D.

Apenas num regime puramente imanentista e materialista, em que não há qualquer forma de transcendência e em que tudo deriva da matéria, pode a inteligência ser considerada *natural*, dado que toda ela, incoativamente, provém da natural materialidade. Em termos, por exemplo, da tradição cristã, uma vez que a inteligência – e tudo o resto – é criada por Deus, não pela natureza (esta, sim, foi criada por Deus), é produto não-natural, mas artificial, produto da arte criadora de Deus.

Talvez o grande equívoco relativo ao que se considera ser a moderna «inteligência artificial» consista em pensar que esta é diversa da preexistente por ser «artificial»; ora, a relação que há entre a criação da inteligência por Deus, mormente a humana, à sua imagem e semelhança, e a criada pelo ser humano, também à sua imagem e semelhança, é analógica, logo, diferente, mas não diversa.

"Talvez o grande equívoco relativo ao que se considera ser a moderna «inteligência artificial» consista em pensar que esta é diversa da preexistente por ser «artificial»".

No entanto, quer no caso materialista-naturalista quer no caso divino-artificialista, a IA é *produto consequente* do que é o acto do criador e do que é o criador enquanto acto; prolonga e reflecte sempre o criador.

Em que é que tudo isto diz respeito à ética, como aqui definida? Em que *isso que é a IA*, seja a moderna de humana criação, segundo um modo materialista sequencial e evolutivo, seja a divina, *depende sempre de isso que a criou*.

Pode pensar-se que o cerne de eventuais questões de correcção ou de perversidade do uso de tais inteligências depende fundamentalmente de isso que as vai usar/ser em acto. Todavia, qualquer uso possível e qualquer acto possível, como actualização de tal possibilidade, depende fundamentalmente da estrutura lógica da possibilidade da inteligência, do seu exercício, na relação com a estrutura lógica da possibilidade de escolha.

Se o algoritmo de possibilidade de intelecção determinar um certo horizonte de possibilidade, então, todo o acto de inteligibilidade só se poderá exercer no interior – lógico – das possibilidades postas por tal algoritmo: a inteligência humana pode intuir a existência de realidades metafísicas infinitas e em número infinito – por exemplo, algumas de ordem matemática –, mas não as domina no que são enquanto realidades infinitas em acto, por exemplo, o exacto número que corresponde a «raiz quadrada de dois».

D.

O que se disse acerca deste tipo de realidade pode ser dito para a IA, sujeita, sempre, aos constrangimentos lógicos (que são ontológicos) dos algoritmos a que obedece, mesmo que tais algoritmos sejam algo como auto-evolutivos.

Primeira conclusão: a acção que deriva da actualização das possibilidades de uso da IA e as suas consequências remetem a responsabilidade, em última análise, para *quem criou tal IA*, não apenas por tal criação, mas também por todos os algoritmos decorrentes criados e a criar, seu poder lógico, sua capacidade consequencial. Tal aplica-se mesmo ao caso divino, em que o modo como o divino criou a inteligência humana responsabiliza tal divino. Esta conclusão é inescapável.

"(...) a acção que deriva da actualização das possibilidades de uso da IA e as suas consequências remetem a responsabilidade, em última análise, para quem criou tal IA".

Segunda conclusão, todo o uso que se fizer da IA, independentemente do que ficou estabelecido na primeira conclusão, responsabiliza o utilizador, seja este utilizador criador de algoritmos, seja apenas mero utilizador de tais algoritmos.

Terceira conclusão: a responsabilidade não é inter-transferível: criador e utilizador são ambos responsáveis; ambos devem ser responsabilizados, para o bem e para o mal.

Em si mesma, a pura possibilidade da IA, como a pura possibilidade da inteligência humana 'de carne' – em redução materialista ou de origem divina –, são isso mesmo: meras possibilidades.

Qualquer possibilidade de acção, isto é, ética, é, na realidade, ontológica, pois implica com seres. Ora, é no fruto da escolha que reside a eventual problematicidade: *que realidade emerge do uso da inteligência*, qualquer seja?

Como com a inteligência humana incarnada, com a moderna IA, se se contribuir para o real bem-comum humano – em regime de «aldeia global» –, *óptimo*; se se contribuir para algo diverso de tal bem-comum, por exemplo, para o bem de um tirano, de uma oligarquia ou mesmo de uma maioria, mas sacrificando qualquer minoria, então, *péssimo*, pois apenas se estará mantendo o habitual regime de uso da inteligência e da vontade – do que se intui e do que se escolhe –, de modo que parte dos seres humanos se engrandeça às custas de outra parte dos seres humanos.

"(...) com a moderna IA, se se contribuir para o real bem-comum humano – em regime de «aldeia global» –, óptimo; se se contribuir para algo diverso de tal bem-comum (...) sacrificando qualquer minoria, então, péssimo".

D.

Pensamos que o modo como Stephen Spielberg põe a questão ética, no filme cujo argumento também escreveu e que tem o nome de IA, segundo o qual a humanidade, por iniquidade anti-humana (e anti-mecânica), se auto-aniquila, sendo substituída por algoritmos que evoluíram – porque partiram de um padrão de bem-querer logificado algoritmicamente –, merece ser meditado, ainda que como mito ou metáfora que é, tendo em conta o grande risco ético e político da humanidade que convive com a IA, risco que existe não por 'culpa' da IA, mas por culpa da humanidade, ultrapassável por um algoritmo evolutivo que poderá acabar por a substituir.

"Pensamos que o modo como Stephen Spielberg põe a questão ética (...) merece ser meditado, ainda que como mito ou metáfora que é, tendo em conta o grande risco ético e político da humanidade que convive com a IA (...)."

Será que o futuro da humanidade será apenas lógico, na forma de algoritmos «melhor-que-humanos»?

Américo José Pereira

Professor Auxiliar da Faculdade de Ciências Humanas
(Católica em Lisboa)

D.

INTELIGÊNCIA ARTIFICIAL E ÉTICA

Relacionando com outras máquinas que substituem o trabalho muscular, a Inteligência Artificial atua no âmbito das atividades cognitivas. A Inteligência Artificia generativa mostra capacidades e realiza tarefas correntemente associadas à inteligência humana, como se fossem executadas por pessoas. Aprende a linguagem humana nas diferentes expressões, e utiliza-a com destreza; pode tomar decisões para as pessoas, sobre elas e com elas. Ou seja, à máquina são atribuídas capacidades e tarefas habitualmente consideradas propriamente humanas, e realizadas com maior eficácia. Dispõe de uma mole infinita de dados, organizados de modo a que o utente, a uma pergunta colocada, receba a resposta mais premente, com rapidez e precisão muito maior do que mente humana.

"A Inteligência Artificia generativa mostra capacidades e realiza tarefas correntemente associadas à inteligência humana, como se fossem executadas por pessoas."

Está presente em muitos aspetos da vida quotidiana, a nível pessoal e a nível social; ajuda em multíplices tarefas. Transforma e continuará a transformar a economia, a segurança, o trabalho, o emprego, a vida das pessoas. Não é agora ocasião para elencar as múltiplas vantagens da Inteligência Artificial, em diferentes âmbitos. Também são conhecidos os riscos, sobretudo a nível da determinação dos comportamentos das pessoas, na difusão de notícias, imagens e sons que parecem verosímeis, mas são falsos.

D.

Não é por acaso que várias instituições a nível internacional têm mostrado preocupação e proposto regulamentação. É destes dias a aprovação da regulamentação de Inteligência Artificial pelos 27 países da União Europeia. Este o nível jurídico. Fixemos-mos nas dimensões antropológica e ética.

Não se trata de uma realidade neutra, de uma ferramenta inerte, sem ação própria. Pode influenciar e modificar, mesmo em profundidade a realidade da pessoa humana, no seu sentido e significado e na sua complexidade. A questão é: A Inteligência Artificial potencia ou substitui a mente humana? É a pessoa que a comanda ou é a máquina que determina? Pode considerar-se uma revolução antropológica; juntamente com esta coloca-se a questão da ética.

"A questão é: A Inteligência Artificial potencia ou substitui a mente humana?"

É comum e adquirida a denominação "Inteligência Artificial". Vale a pena determo-nos rapidamente sobre esta designação, e o sentido analógico que tem. Comecemos pelo adjetivo; "artificial" contrapõem-se a "natural", ou seja, a inteligência humana. "Artificial" quer dizer "feita por arte, por engenho"; algo construído pela inteligência humana. Vejamos o substantivo, "inteligência", na origem latina o vocábulo é formado por *intus* (dentro) e *leggere* (recolher, reunir, ler). literalmente ler dentro, ver em profundidade, perceber o significado profundo das coisas.

A Inteligência Artificial não faz isso. Tem uma capacidade imensamente maior do que o ser humano, de memorizar, de processar uma mole quase infinita de dados já catalogados, pondo-os à sua disposição, depois de perguntas ou de interesses. Lê-os, faz combinações e elabora respostas; atua numa dimensão lógica, racional, padronizada. Mas não decifra o significado e os sentimentos humanos; isso compete apenas à pessoa.

Por outro lado, é necessário ter bem presente na dimensão humana, a instância tão valorizada a que se chama "Inteligência Emocional". É relevante na vida, na determinação dos valores na concretude das situações e na decisão moral; é absolutamente pessoal, não uniformizada entre as pessoas e até mesmo na mesma pessoa, variando conforme as condições e as circunstâncias.

No mundo da racionalidade lógica em que a inteligência artificial se move, é iluminador ter presente a frase de Pascal, filósofo e matemático: "o coração tem razões que a razão desconhece". "Coração" aqui pode traduzir-se por "emoções", "sentimentos", "intuições"; realidades que estão para lá da pura lógica, e que indicam o sentir profundo da pessoa singular, tantas vezes na origem da opção moral concreta, e que parece ter pouca "lógica". O papa Francisco, na mensagem para o Dia Mundial das Comunicações Sociais, 2024, com o título "Inteligência Artificial e sabedoria do coração: para uma comunicação plenamente humana", enfatiza esta vertente.

D.

“No mundo da racionalidade lógica em que a inteligência artificial se move, é iluminador ter presente a frase de Pascal, filósofo e matemático: “o coração tem razões que a razão desconhece”.”

“Neste tempo que corre o risco de ser rico em técnica e pobre em humanidade, a nossa reflexão só pode partir do coração humano. Somente dotando-nos dum olhar espiritual, apenas recuperando uma sabedoria do coração é que poderemos ler e interpretar a novidade do nosso tempo e descobrir o caminho para uma comunicação plenamente humana. O coração, entendido biblicamente como sede da liberdade e das decisões mais importantes da vida, é símbolo de integridade e de unidade, mas evoca também os afetos, os desejos, os sonhos, e sobretudo é o lugar interior do encontro com Deus. Por isso a sabedoria do coração é a virtude que nos permite combinar o todo com as partes, as decisões com as suas consequências, as grandezas com as fragilidades, o passado com o futuro, o eu com o nós. Esta sabedoria do coração [...], compreender as interligações, as situações, os acontecimentos e descobrir o seu sentido. Sem esta sabedoria, a existência torna-se insípida”.

Já referimos a dimensão ética. Nos alertas que se dão, nas reflexões que se fazem, na regulamentação que se pede, nas propostas e projetos que se elaboram em diferentes instâncias, a dimensão ética está sempre presente; todos a referem. Parece-nos que com diferentes entendimentos.

Entendemo-la não como ética aplicada, isto é, elaboração e enumeração de princípios éticos gerais, que se apliquem a realidades concretas particulares, no caso a inteligência artificial, para poder ter confiança que atua para servir valores humanos. Propomos a sua compreensão como afirmação de valores e busca de soluções para problemas normativos com objetivos práticos, ao mesmo tempo que, em circularidade hermenêutica, coloca sempre novas questões, de modo personalizado e em novo contexto, pois a vida pessoal tem unidade, não homogénea e unívoca, mas poliédrica.

“Entendemo-la não como ética aplicada, isto é, elaboração e enumeração de princípios éticos gerais, que se apliquem a realidades concretas particulares, no caso a inteligência artificial, para poder ter confiança que atua para servir valores humanos.”

Isto realiza-se na e pela consciência pessoal, ou melhor, autoconsciência que a máquina não tem. Detenhamo-nos um pouco. É centro e central no âmbito ético e moral; e pessoal e intransmissível, a primeira e última da vida moral. Conjuga-se com os conceitos de “inteligência humana” e de “inteligência emocional”.

D.

Tem duas dimensões interligadas, a psicológica e a moral. A primeira refere-se à capacidade de apreender, conhecer o que está em jogo, propor valores éticos a realizar. A segunda tem uma função judicativa, isto é, elaborar juízos morais sobre as ações ou omissões tendo em atenção três fatores a ponderar em simultâneo: o assunto de que se trata, e a sua relação maior ou menor com os valores em causa, isto é, a dimensão objetiva; a intenção do agente, ou seja, a dimensão subjetiva, e as várias circunstâncias que podem ainda ser mutáveis.

Nas problemáticas morais, tantas vezes, o que a lógica estrita indica como solução, não deve nem pode ser realizado; a razão do coração viu, sentiu, intuiu e impeliu em sentido diferente, mais correto. Trata-se do discernimento, entendido como exercício do juízo moral. De outro modo haveria uma resposta padronizada, impessoal e, portanto, irresponsável.

Parece-nos que a Inteligência Artificial pode ajudar na dimensão psicológica, isto é, no conhecimento a partir dos dados que lhe são fornecidos, não na dimensão moral, porque não tem nem inteligência em sentido próprio, nem coração. Carece de ser emocional, não tem liberdade nem responsabilidade. Afinal, é uma inteligência programada, lógica, sem autoconsciência, porque é produzida como imitação; não está viva, não respira, não tem metabolismo; é uma máquina que pode ser ligada e desligada.

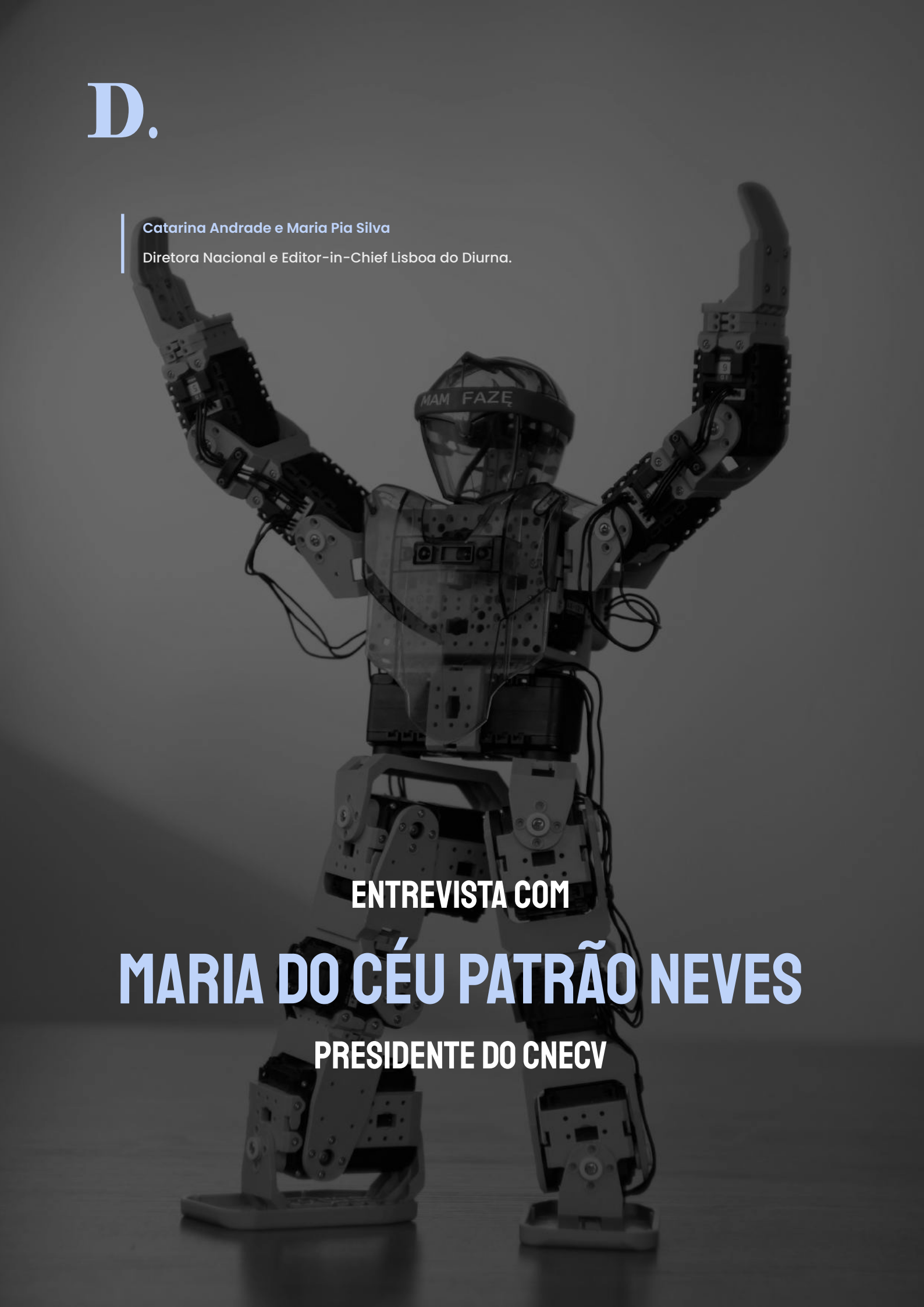
Jerónimo Trigo

Professor Auxiliar da Faculdade de Teologia (Católica de Lisboa)

D.

Catarina Andrade e Maria Pia Silva

Diretora Nacional e Editor-in-Chief Lisboa do Diurna.



ENTREVISTA COM

MARIA DO CÉU PATRÃO NEVES

PRESIDENTE DO CNECV

D.

Maria do Céu Patrão Neves

Maria do Céu Patrão Neves é Professora Catedrática de Ética, Presidente do Conselho Nacional de Ética para as Ciências da Vida e Vice-presidente do Grupo Europeu de Ética. Numa tarde de inverno, conversou connosco sobre os desafios éticos que nos reserva a rápida evolução da Inteligência Artificial, numa reflexão sobre as implicações da mesma na medicina, no merca-do de trabalho e na educação.

Confidenciou-nos que se hoje ganhasse a lotaria, amanhã estaria a fazer exatamente a mesma coisa.

Numa edição em que só se escreve sobre máquinas, tivemos a sorte de poder ouvir uma voz tão humana como a sua.

Se não tivermos programas adequados a todos os segmentos da população, vamos deixar pessoas para trás. E uma sociedade que se permite abandonar ou negligenciar alguns cidadãos é uma sociedade injusta e não solidária.

Como é que definiria a relação entre a inteligência artificial e a ética?

Na União Europeia toda a regulamentação da inteligência artificial se fundamenta em valores e princípios éticos, sendo que a União Europeia é o primeiro bloco geopolítico a proceder a uma regulamentação robusta e, simultaneamente, flexível da IA na sua abordagem através de níveis de risco. Os antecedentes mais próximos datam de 2018 e a Comissão Europeia fez questão de começar precisamente pela reflexão ética e só depois elaborar aquilo que são normativas jurídicas, na observância estrita dos valores éticos. Por outro lado, há também (e importa dizê-lo) uma corrente bastante forte, sobretudo entre engenheiros e empresários, que ativamente reduz a inteligência artificial a uma designada ferramenta, isto é, a um objeto que é verdadeiramente inerte, axiologicamente neutro, que remete a agência para o utilizador e, assim, desvaloriza ou minimiza também os seus reais e potenciais impactos sociais. O objetivo é claro: anular quaisquer *baias* ao progresso da inteligência artificial. Ora, as *baias, ou linhas condutoras, éticas e jurídicas* são indispensáveis para garantir que o progresso tecnológico não serve apenas o interesse de poucos, mas antes beneficia efetivamente a humanidade.

D.

"(...) estas baias (...) são indispensáveis para garantir que o progresso tecnológico não serve apenas o interesse de poucos, mas antes beneficia efetivamente a humanidade."

Quais é que são os princípios éticos universais aplicáveis à Inteligência Artificial?

Os princípios éticos que estão plasmados na Declaração Universal dos Direitos Humanos, nas suas três gerações (civis e políticos; sociais, económicos e culturais; e de titularidade coletiva) aplicam-se, obviamente, a todo e qualquer uso da Inteligência Artificial. Destaco, em todo o caso, no plano individual, a dignidade humana, no plano coletivo, a justiça social. Princípios como a não discriminação e a privacidade também têm particular importância na utilização da Inteligência Artificial e são princípios que esta, por vezes, tem colocado em causa. Outros ainda bastante determinantes serão: o da transparência, o da explicabilidade e o da responsabilidade.

Estamos a falar de princípios que são, como dita a própria noção, mandados de otimização. Nem sempre deve ser fácil perceber em que medida é que devem ser aplicados.

Estes princípios devem ser sempre respeitados. O problema é que, à medida que a Inteligência Artificial avança, os próprios responsáveis pelo seu desenho e processamento reconhecem existirem situações em que introduzem dados no sistema, recolhem os resultados, mas a explicabilidade do processo já não é absoluta, e, por isso, a transparência também não o é. E este é o problema ético grave!

Qual é que é então a posição do Conselho Nacional de Ética para as Ciências da Vida em relação aos avanços tecnológicos e ao papel da ética nestes mesmos avanços?

Em termos gerais, o Conselho aprecia cada inovação tecnológica no seu real e potencial impacto, nas pessoas, na sociedade, tomando como critério precisamente os valores e princípios éticos identitários da humanidade. Tem-se sempre a preocupação de não destabilizar o progresso cientifico-tecnológico, mas antes de o orientar, no sentido de máximo benefício para a sociedade, para a humanidade, evitando e reprovando que o avanço cientifico-tecnológico possa ficar cativo e refém de interesses sectários (como, por exemplo, académico-científicos, económico-financeiros ou geopolíticos).

Nos filmes de ficção científica, costumamos ver robots que vencem a humanidade e dominam o mundo. Considera que vai haver uma guerra de robots contra humanos, ou, na realidade, a guerra vai fazer-se entre seres humanos?

D.

Eu diria que a ficção científica projeta receios e pode prevenir desgraças, merecendo a nossa atenção. Em todo o caso, e lamentavelmente, as guerras entre humanos têm sido uma constante na história da humanidade, independentemente da inteligência artificial, e com dramáticos momentos de forte recrudescimento como aquele que estamos a viver hoje. Se queremos associar a guerra entre humanos à utilização da Inteligência Artificial, diria que esta é uma realidade que já está presente na Europa e nós tanto podemos apontar o seu contributo para a eficiência do ato da guerra (por exemplo, na destruição do que é considerado um alvo vital sem atingir civis), como também podemos apontar o distanciamento que cria entre o agressor e a vítima, anonimizando as duas e convertendo a violência numa ação puramente técnica, axiologicamente neutra e, assim, anulando também a perceção do sofrimento e a possibilidade da empatia. Estes aspetos são talvez os mais importantes e não foi por acaso que o Santo Padre dedicou a sua mensagem deste ano, no Dia Mundial da Paz, à paz e à Inteligência Artificial.

"Eu diria que a ficção científica projeta receios e pode prevenir desgraças (...)."

E em relação à "guerra" que se cria por força das desigualdades económicas e das diferentes velocidades de desenvolvimento?

O desenvolvimento tecnológico não é igualitário e temos clara consciência de que há blocos geopolíticos que têm uma civilização tecnológica muito mais avançada do que outros. Neste contexto, impera o princípio da solidariedade, estabelecido na Declaração Universal de Bioética dos Direitos Humanos, de 2005, e que preconiza a partilha dos benefícios cientifico-tecnológicos com os países menos avançados. Mas, naturalmente, estou a falar de benefícios para a subsistência, a saúde, o bem-estar, a realização pessoal, a justiça entre os povos do mundo.

Na era da recolha massiva de dados, como é que podemos compatibilizar a necessidade de avanço tecnológico com a preservação da privacidade individual?

É já um lugar-comum dizer que a literacia digital é absolutamente essencial para uma boa utilização dos recursos digitais. Eu gostaria de acrescentar uma chamada de atenção para o facto de a atual ilusão da privacidade nas nossas sociedades ser também da responsabilidade do próprio cidadão, isto é, para além de uma ganância de dados por parte de empresas e de instituições diversas, são muitas vezes os próprios cidadãos que facilmente cedem os seus dados em troca de muito pouco ou mesmo nada. A promessa de uma loja de enviar um talão oferta no dia dos anos do cliente leva-o a dar muitos mais dados do que apenas a data de nascimento. Por outro lado, basta entrarmos nas redes sociais para encontrarmos aí um verdadeiro palco, um verdadeiro mercado aberto de dados oferecidos pelos cidadãos sobre si próprios e, às vezes, também sobre outrem. Este aspeto tem de ser sublinhado fortemente para que cada um de nós seja também o verdadeiro guardião dos seus dados.

No desenvolvimento de mecanismos de Inteligência Artificial, podem ser incorporadas considerações éticas desde o início do processo de criação. Quais é que são as melhores práticas para garantir um desenvolvimento ético da inteligência artificial?

Ser desenhada para servir as finalidades humanas, no respeito pelos princípios identitários da humanidade, ser explicável ao longo de todo o processo e manter-se sob controle humano em todas as suas utilizações. Diria que são estes os princípios chave.

Tendo em conta que a regulamentação desempenha um papel fundamental na orientação ética da Inteligência Artificial, quais é que são os benefícios e desafios associados à criação de normas e regulamentos para disciplinar o desenvolvimento e utilização da Inteligência Artificial?

O primeiro benefício da regulação é estabelecer parâmetros e aceitabilidade da própria Inteligência Artificial. Paralelamente, o primeiro desafio é o da regulação ser observada, ou seja, o seu cumprimento deve ser efetivo e ser supervisionado. A este nível, podemos acrescentar que hoje projetos científicos que extravasam a regulamentação deixam de poder ser financiados pela Comissão Europeia. Isto é importante porque frequentemente formulamos regras que depois não são cumpridas e não há consequências. Esta é a a realidade, mas, por outro lado, estão também a ser criados mecanismos para penalização do não cumprimento e de incentivo ao cumprimento. Um segundo benefício da regulamentação é contribuir para uma cultura de respeito para potencializar o maior benefício para maior parte das pessoas. Refiro-me a uma cultura, a uma prática interiorizada, assimilada e comummente seguida, e não apenas de uma normativa exterior a cumprir, assim gerando um ambiente em que nos sintamos tranquilos e em que possamos contribuir para que se torne o mais favorável para a humanidade. O AI Act, aprovado por unanimidade em reunião de embaixadores da passada semana, prevendo-se que haja uma pontuação favorável no Conselho Europeu no próximo mês de abril, será a primeira regula-mentação robusta no mundo em Inteligência Artificial e a expectativa é que possa constituir um verdadeiro paradigma para que o seu modelo venha a ser adaptado por outras outras áreas geopolíticas.

"O primeiro benefício da regulação é estabelecer parâmetros e aceitabilidade da própria Inteligência Artificial. Paralelamente, o primeiro desafio é o da regulação ser observada, ou seja, o seu cumprimento deve ser efetivo e ser supervisionado."

Dada a crescente automação, como é que podemos lidar com as decisões tomadas por máquinas, especialmente quando se trata de questões ligados à saude? E quais é que são os princípios éticos de regulamentação e utilização desses mesmos sistemas?

Os princípios éticos antes referidos mantêm-se como obrigatórios na atividade clínica, podendo acrescentar-se alguns identitários dos cuidados da saúde - como a beneficência do doente, a sua autonomia, a equidade no acesso da distribuição de cuidados e bens de saúde, ou a atenção à vulnerabilidade física, emocional e social que impõe sempre uma acrescida e proporcional proteção dessa mesma pessoa. Neste contexto, diria ser absolutamente necessário é que a decisão final seja humana, isto é, seja do médico ou da equipa de saúde. Só assim se garante a perspetivação do doente como um todo, uma unidade singular, não numeralizável, não objetivável e não padronizável, se promove o tratar da pessoa doente e não da doença. As doenças não existem em abstrato, mas nas pessoas, sendo que a mesma doença em pessoas distintas pode ter manifestações diferentes também. Se um assistente digital procede a um diagnóstico, traça uma terapêutica e toma uma decisão clínica, como se pode exercer o direito do doente a uma segunda opinião (quando faz parte dos direitos humanos ter uma segunda opinião sobre a sua realidade clínica)? Afinal, qualquer assistente digital, vai dizer exatamente o mesmo...E se um médico proceder a uma diferente avaliação, contrariando o assistente digital? Ele será certamente responsável pelo ato clínico? Mas o médico também pode ter uma diferente perceção do estado clínico do doente e optar pela recomendação do assistente digital...Qual a responsabilidade atribuível ao assistente digital? São exemplos da complexidade das questões sobre as quais estamos a refletir.

"(...) é absolutamente necessário é que a decisão final seja humana (...) na perspetivação do doente como um todo, no tratar da pessoa doente e não da doença."

Já referimos também a transparência como princípio fundamental da ética, crucial para fundamentar a confiança da sociedade em relação à Inteligência Artificial. Como é que podemos garantir essa transparência nas decisões algorítmicas e estabelecer mecanismos eficazes de responsabilização?

A transparência obtém-se e promove-se a partir do conhecimento dos processos por que os sistemas operam, o que, por seu turno, vai permitir também um melhor apuramento das responsabilidades. As duas questões estão interligadas. O progresso da Inteligência Artificial tende, não obstante, a diluir a responsabilidade num grande domínio de difusa ambiguidade. Esta é a perceção comum que terá de ser necessariamente contrariada, porque se não houver capacidade técnica e normativas jurídicas para uma justa atribuição das responsabilidades, é a vítima, o lesado, que se deixa desprotegido.

D.

Nesse sentido, a Inteligência Artificial poderá então ser implementada para melhorar quer a precisão, quer a rapidez no diagnóstico médico, porque caso contrário não seria implementada. Quais é que são depois os desafios éticos associados à dependência de algoritmos para os diagnósticos?

É como diz. A inteligência artificial não seria utilizada se não trouxesse grandes vantagens. Trata-se de um truísmo, mas vale a pena apontá-lo: não estaríamos aqui a discutir a inteligência artificial se não houvesse vantagens óbvias na sua utilização. E, efetivamente, no domínio da Medicina tem havido enormes progressos. Eu aproveitaria para referir que no relatório do CNECV de 2022 há uma reflexão sobre cinco grandes domínios em que a Inteligência Artificial tem trazido ganhos em saúde, a saber: o da investigação biomédica, da assistência clínica, da gestão hospitalar, da administração da saúde, o da educação para profissionais de saúde (formação) e para a população (literacia). Os benefícios são inquivocamente amplos. Respondendo mais diretamente à questão, isto é, no que se refere às vantagens de utilização da Inteligência Artificial no diagnóstico, este ganha em rapidez e elevado grau de precisão para várias patologias, e muito em particular em doenças raras, o que se traduz em ganhos para a saúde, especialmente para o doente, mas também na otimização dos recursos humanos, equipamentos, financeiros.

"(...) nós não estaríamos aqui a discutir a inteligência artificial se não houvesse vantagens óbvias na sua utilização. E, efetivamente, no domínio da Medicina tem havido enormes progressos."

Gostaria também de sublinhar a importância da validação prévia destas ferramentas utilizadas em saúde e da indispensável competência do utilizador para o seu uso, nomeadamente na introdução de dados e interpretação dos mesmos, não deixando de referir também a necessária qualidade dos dados. Será uma falácia pensar que bastará ter um software muito preciso para termos resultados proporcionalmente rigorosos. É preciso validar o instrumento, confirmar a competência do utilizador e assegurar a qualidade dos dados. Sem esta tríade não temos garantias das efetivas vantagens da utilização da Inteligência Artificial.

"É preciso validar o instrumento, confirmar a competência do utilizador e assegurar a qualidade dos dados. Sem esta tríade, não temos garantias das efetivas vantagens da utilização da Inteligência Artificial."

As cirurgias realizadas e assistidas por robots e mecanismos de Inteligência Artificial também são, cada vez mais, uma realidade próxima. Considera que é eticamente plausível que sistemas de Inteligência Artificial possam realizar cirurgias? O que é que tem a ética a dizer em relação à segurança do paciente e à responsabilidade médica durante procedimentos cirúrgicos automatizados?

D.

Os robots cirúrgicos podem ter um desempenho de alta precisão e ser também impermeáveis a quaisquer fatores humanos que possam, eventualmente, perturbar um desempenho rigoroso. São aspetos que claramente jogam a favor da Inteligência Artificial e é neste sentido que podem ser, ou já são, um recurso de elevado valor no bloco operatório. Sublinho, de novo, a exigência de validação prévia destas ferramentas de intervenção e da formação adequada dos operadores desses recursos, porque estando estes dois aspetos assegurados, eu diria que temos salvaguardada a segurança do paciente e a assunção da responsabilidade do médico. Temos de apreciar a utilização da Inteligência Artificial não apenas de uma forma simplificada, mas na complexidade do seu desempenho.

Podendo a Inteligência Artificial sugerir opções de tratamento com base em exames médicos, como é que podemos garantir a tomada de decisões clínicas de forma ética, que tenha em linha de conta o expertise dos seres humanos, por um lado, e a autonomia dos pacientes, por outro?

A competência especializada do médico é preservada e o paciente pode beneficiar dela ao mesmo tempo que, enquanto parceiro da relação terapêutica, exerce o seu direito à autonomia, se a decisão for humana, se a decisão for conjuntamente do médico e do paciente. A Inteligência Artificial deve intervir como coadjuvante, como complementar, e não como substituto do conhecimento e experiências especializadas do profissional de saúde ou da decisão livre do doente. Desta forma, mesmo que o projeto terapêutico tenha sido delineado conjuntamente pela inteligência artificial e pelo médico assistente, o doente terá sempre oportunidade de melhor conhecer a sua realidade clínica e de tomar uma decisão informada acerca do que considera desejável para si.

"A competência especializada do médico é preservada e o paciente pode beneficiar dela ao mesmo tempo que, enquanto parceiro da relação terpêutica, exerce o seu direito à autonomia, se a decisão for humana, se a decisão for conjuntamente do médico e do paciente."

D.

MERCADO DE TRABALHO

A Inteligência Artificial pode vir alterar o mercado de trabalho em diversos setores, podendo mesmo vir a extinguir postos de trabalho. Como é que podemos garantir que essas mudanças sejam éticas e tenham em conta o impacto social, especialmente em relação à segurança no emprego?

A Inteligência Artificial já está, há alguns anos, a alterar profundamente o mercado de trabalho, sendo cada vez mais numerosas as atividades em que o ser humano pode ser substituído por sistemas inteligentes com uma eficácia sempre crescente. Lembrar-se-ão da inauguração da primeira Web Summit em Portugal, em 2017, em que a robô humanóide Sophia dizia: "Vamos tirar-vos os empregos!". Não sei se era ficção científica, se era uma promessa, se era uma ameaça... Interessa-me, sobretudo, sublinhar o facto de segmentos da população, profissionais, que têm benefícios diretos do progresso da Inteligência Artificial frequentemente afirmarem que são apenas os trabalhos rotineiros e aborrecidos que irão ser substituídos. Hoje, sabemos que trabalhos que exigem elevado conhecimento e competência técnica podem ser substituídos pela Inteligência Artificial em prol da eficácia do desempenho. Nós não temos sistemas inteligentes apenas em fábricas ou na indústria, mas também na prestação de cuidados de saúde, na advocacia, no ensino, na gestão de pessoal, nas finanças, etc..

"(Na) inauguração da primeira Web Summit em Portugal a robô humanóide Sophia dizia: "Vamos tirar-vos os empregos!". Não sei se era ficção científica, se era uma promessa, se era uma ameaça..."

O outro argumento que importa desmistificar é o de que sempre houve empregos que desapareceram e outros que surgiram. É uma verdade ao longo da história da humanidade. O fator inédito hoje é a rapidez desta dinâmica. Se, no passado, era ao longo de décadas que uma profissão ia enfraquecendo e outras se iam avolumando, agora é no espaço de meses que qlgumas se extinguem e outras se impõe com uma necessidade e urgência prementes. A capacidade de adaptação a esta tão célere mudança laboral é extraordinariamente exigente. Há segmentos da população que não têm capacidade para uma adaptação tão rápida e tão profunda, pelo que são ultrapassados e abandonados pelo progresso...

O terceiro aspeto, o terceiro mito que aqui gostaria de denunciar é o de que, com a Inteligência Artificial, vamos todos ganhar muito mais tempo de lazer, de que podemos, afinal, não investir muito no trabalho. Ora, o mundo está organizado de tal forma que não é possível o financiamento para as nossas necessidades básicas sem o mercado de trabalho. Além disso, o trabalho também não é apenas uma ocupação obrigatória. É um espaço de realização pessoal (para mim é um espaço de realização pessoal e eu costumo dizer que se me saísse a lotaria (eu não jogo), no dia seguinte iria fazer exatamente a mesma coisa, trabalhar!).

Como é que podemos mitigar as disparidades de oportunidades de emprego geradas pela implementação de tecnologias inteligentes?

Possivelmente antecipando de forma realista os impactos negativos e projetando medidas mitigadoras desses mesmos impactos. Neste âmbito, diria que importa disponibilizar programas de reconversão laboral para adultos, programas formativos para jovens e de literacia digital para todos.

Podemos evitar a discriminação algorítmica no processo de contratação e de promoção?

A utilização de sistemas de Inteligência Artificial para analisar currículos, para selecionar candidatos para diferentes empregos, para avaliar níveis de produtividade, dispensar trabalhadores e decidir promoções, etc... está-se a tornar cada vez mais comum. Reconheço que, num primeiro momento, se existem, por exemplo, 100 pessoas para dois empregos, se recorra a sistemas de Inteligência Artificial para verificar se todos os processos cumprem os requisitos solicitados. Este primeiro rastreio pode ser feito com enormes vantagens, por sistemas de Inteligência Artificial. Porém, prolongar o processo de seleção recorrendo apenas à Inteligência Artificial é descurar, ignorar, dimensões do humano que não são suscetíveis de ser quantificadas, objetiváveis, o que é igualmente válido quer para a obtenção de emprego, quer para a atenção da promoção, quer para o despedimento, o que, neste último caso, sabemos estar a acontecer no presente em grandes empresas, nomeadamente Big Tech. É reduzir o ser humano ao que é quantificável, empobrece-lo ao descartar a singularidade de cada pessoa... e nem sequer é um tratamento justo. Por isso, do ponto de vista ético, eu diria que é altamente reprovável.

"É reduzir o ser humano ao que é quantificável, empobrece-lo ao descartar a singularidade de cada pessoa...e nem sequer é um tratamento justo."

Quais é que são os esforços éticos para garantir que as decisões de Inteligência Artificial não perpetuam preconceitos existentes no mercado de trabalho?

Os vieses algorítmicos são hoje bem conhecidos, o que acaba por ser um aspeto positivo na medida que, ao serem reconhecidos, haverá também vontade de os contrariar. Podemos contrariá-los construindo equipas para trabalhar em sistemas de Inteligência Artificial que sejam socialmente mais diversas, aumentando o volume de dados recolhidos e ajustando continuamente a configuração dos dados no sentido de ir erodindo os atuais vieses que decorrem em grande parte da própria sociedade em que vivemos. Às vezes, diz-se que estes vieses algorítmicos são um reflexo da nossa sociedade atual, mas são mais do que um mero reflexo; são uma amplificação. É como se os sistemas de inteligência artificial constituíssem uma caixa de ressonância que acaba por aumentar, amplificar as disparidades e as desigualdades que encontramos na sociedade atual.

Como é que a Inteligência Artificial pode ser utilizada para personalizar a experiência dos alunos, olhando às suas necessidades individuais?

Eu insisto que a Inteligência Artificial pode ser um instrumento complementar e não substituto de uma relação pessoal, nomeadamente da relação entre alunos e professores. Na sua função coadjuvante, colaborativa, a Inteligência Artificial pode efetivamente satisfazer necessidades específicas, particulares, pessoais, que devem depois ser integradas no projeto educativo holístico que as próprias instituições de ensino e os seus agentes devem construir e proporcionar aos seus alunos. O aluno não pode ser abandonado à Inteligência Artificial. Deve ser acompanhado, na sua busca autónoma, através da Inteligência Artificial. E esse acompanhamento da sua aquisição de conhecimentos e satisfação de curiosidade intelectual deve ser acompanhado por quem tem outros conhecimentos e outras experiências, e que pode e deve também cultivar o sentido crítico do aluno. Um dos graves problemas nesta matéria é a ideia, lamentavelmente bastante difundida, de que tudo o que corre na internet é verdadeiro. Hoje, esta não é de todo uma realidade. Precisamos de desenvolver um sentido crítico. Se o aluno (quanto mais jovem pior) for deixado sozinho na pesquisa das suas necessidades pessoais – que deve fazer! - é abandonado. E o sistema não o pode abandonar, deve estimular essa autonomia, mas acompanhá-lo, apoiá-lo, sobretudo através do desenvolvimento do projeto holístico de desenvolvimento pessoal e desenvolvimento do próprio sentido crítico.

"O aluno não pode ser abandonado à Inteligência Artificial. Deve ser acompanhado, na sua busca autónoma, através da Inteligência Artificial. E esse acompanhamento da sua aquisição de conhecimentos e satisfação de curiosidade intelectual deve ser acompanhado por quem tem outros conhecimentos e outras experiências, e que pode e deve também cultivar o sentido crítico."

Na realidade, hoje o problema é exatamente o contrário do que existia antes. Ou seja, temos tanta informação que nem sequer sabemos como a tratar.

Exatamente. O excesso de informação! Nós hoje passámos para os antípodas – antes não tínhamos informação suficiente para dissertar sobre uma qualquer matéria; agora temos tanta informação que nos faltam os critérios de seleção e, sobretudo, o espírito crítico para fazer uma leitura lógica, racional, ponderada, prudente, sobre aquilo que nos é dito.

D.

A Inteligência Artificial pode ser utilizada para detetar e apoiar alunos com necessidades especiais? Quais é que são os cuidados éticos a implementar para assegurar a inclusão e a igualdade?

De facto, a Inteligência Artificial pode ter um papel relevante no apoio a alunos com necessidades especiais. Neste âmbito, gostaria de chamar à atenção para uma tendência (que por vezes sinto crescente) de rotular, classificar e categorizar pessoas. Isto pode conduzir a discriminações posteriores e é também frequentemente um tolhe à própria liberdade individual. Por isso, uma chamada de atenção não apenas para as vantagens, mas também para a derrapagem que pode haver nesta utilização. Quanto à identificação de necessidades especiais, sinto que hoje, por vezes, é exacerbada e que permite inclusivamente alguns abusos. Esta identificação deve ter como objetivo ético fundamental proporcionar condições compensatórias e ajustadas às deficiências detetadas em cada aluno para garantir a igualdade de condições a todos. Este é que é o verdadeiro objetivo ético. E se a Inteligência Artificial for utilizada neste contexto é realmente um fator de equidade e de criação de condições de igualdade para todos - e esse é o objetivo ético a preconizar.

Dada a difusão das plataformas de ensino online, em especial, em tempos de COVID-19, como é que a Inteligência Artificial pode ser incorporada para melhorar o ensino à distância? Quais são os desafios éticos associados à equidade no acesso à educação à distância através de plataformas online?

O primeiríssimo aspeto a apontar é a disponibilização igualitária de equipamentos e de cobertura de rede. Nem vale a pena continuar a discussão se estes dois aspetos não estiverem garantidos: equipamentos de acesso e estabilidade e serviço regular de rede de internet. Sem isso, não vale a pena. A pandemia de COVID-19 foi uma realidade que, tendo detetado os problemas que focou, também constituiu uma motivação enorme, um forte motor, de desenvolvimento da digitalização na educação. Passando o período pandémico, o que vemos hoje é uma proliferação de cursos online, também direcionados para tudo o que possam ser as necessidades e preferências dos alunos de todas as idades. Há inúmeras aulas online. Neste contexto, urge uma validação destes cursos, garantir a sua qualidade, criar ou adequar um sistema de acreditação. Sem este acompanhamento, sem esta garantia de qualidade, sem a acreditação destes cursos, é o próprio cidadão, aluno, consumidor, que pode ser levado ao engano, por vezes pagando avultadas somas, sem ter o retorno esperado. Estes são aspetos para os quais tem de haver uma atenção cada vez maior. Não me refiro, obviamente, a restringir a facilidade que há de acesso à formação online, mas à sua validação através de um sistema rigoroso, por uma autoridade nacional ou internacional reconhecida, precisamente como garantia do melhor serviço e, neste caso, da melhor formação para as pessoas que a procuram.

"Há uma necessidade de validação destes cursos, garantir a sua qualidade, criar ou adequar um sistema de acreditação."

D.

Quais são os principais desafios éticos na implementação generalizada de tecnologias de IA na educação? Considera que faria sentido incluir no ensino básico, e até secundário, disciplinas que ensinassem os alunos a trabalhar com mecanismos de Inteligência Artificial e a tratar as informações daí resultantes? Qual é a posição do Conselho Nacional de Ética para as Ciências da Vida?

O CNECV não tem um parecer, uma deliberação específica, sobre esta matéria (o que, aliás, é natural, porque esta matéria é sobretudo competência do Conselho Nacional de Educação). Não obstante, o Conselho Nacional de Ética claramente considera que a promoção de uma boa utilização de todos os recursos tecnológicos é fundamental para o cidadão, para a sociedade e apresenta-se como um verdadeiro imperativo ético incontornável. Os recursos tecnológicos devem ser pertença, devem ser partilhados por toda a sociedade, devem ser acessíveis a todos os cidadãos, no estabelecimento da máxima igualdade entre todos. Se tal não acontecer, as tecnologias tornam-se fator de discriminação e de injustiça social. Por isso, a única forma de as novas tecnologias não contribuirem para o agravamento da injustiça, a única forma de se perfilarem como um fator de promoção individual e social, é precisamente garantir a educação na sua utilização a todos os cidadãos, independentemente da idade. Não me refiro apenas a programas de ensino para crianças, jovens e jovens adultos, mas programas diferenciados para diferentes níveis etários – não deixar ninguém para trás. Se não tivermos programas adequados a todos os segmentos, vamos deixar pessoas para trás. E a sociedade que se permite abandonar alguns dos seus é uma sociedade injusta e não solidária.

"Se não tivermos programas adequados a todos os segmentos, vamos deixar pessoas para trás. E a sociedade que se permite abandonar alguns dos seus é uma sociedade injusta e não solidária."

Catarina Andrade e Maria Pia Silva

Diretora Nacional e Editor-in-Chief Lisboa do Diurna.

D.

O UNCANNY VALLEY

E O LADO PERTURBADOR DA IA

Atualmente, com a crescente discussão sobre inteligência artificial (IA) e chatbots, somos confrontados com um conjunto de termos em inglês que convém descortinar. Um desses termos é o de "uncanny valley". O termo aparece frequentemente associado ao uso de chatbots, que se têm tornado comuns no relacionamento de algumas marcas com os seus consumidores, designadamente na prestação de serviços de apoio, quer antes da venda ter lugar, quer, durante, quer, sobretudo, nos serviços de apoio ao consumidor no pós-venda. É por isso importante conhecer a que nos referimos quando usamos estes canais conversacionais e quando na literatura da área se usa este termo do "vale do desconhecido".

A utilização do meio digital como um novo canal de marketing, seja para comunicação, venda ou distribuição, proporciona às empresas diversas oportunidades únicas para interagir com seus clientes, tanto atuais quanto potenciais. E um desses canais em ascensão é o que envolve o uso de chatbots, que se têm tornado cada vez mais presentes nas redes sociais, em páginas web e em aplicações de mensagens. Estes programas simulam conversas humanas, atuando como assistentes virtuais para tarefas simples, como oferecer opções de pacotes de acesso à internet, ou outras mais complexas, como realizar compras, prestar aconselhamento financeiro, ou auxiliar numa reclamação ou devolução.

D.

Porém, se o emprego de tecnologias baseadas em inteligência artificial (IA), como os chatbots, está a receber uma atenção crescente por parte do público, também é verdade que, nalguns casos, o ceticismo impera, sobretudo dadas as limitações a que ainda assistimos por parte destes serviços, assim como pelas limitações decorrentes da fraca proteção de dados providenciados. Prevê-se, todavia, que nos próximos dez anos, a incorporação de IA pode complementar e até substituir parcialmente os seres humanos em situações de compra e serviços de atendimento. O rápido crescimento do mercado de chatbots, projetado para atingir 1,34 trilião de dólares até 2024, respalda essas expectativas. No entanto, não são apenas vantagens, e um dos principais desafios é conhecido como "uncanny valley", conforme veremos a seguir.

"(...) a incorporação de IA pode complementar e até substituir parcialmente os seres humanos em situações de compra e serviços de atendimento."

De facto, um problema decorrente do uso de chatbots é que, num futuro próximo, estes programas podem ser confundidos com seres humanos, especialmente se a sua natureza robótica não for revelada. Isto levanta questões éticas envolvendo os clientes, defendendo-se por isso a ideia de que as pessoas devem estar cientes da verdadeira natureza dos chatbots e das implicações decorrentes da interação com um robô em vez de uma pessoa. Para além das questões éticas, e apesar dos benefícios potenciais oferecidos por estes agentes conversacionais, as empresas enfrentam um outro desafio: o da resistência dos clientes em se sentirem à vontade ao lidar com um sistema de IA. De facto, muitas pessoas ainda se sentem desconfortáveis ao interagir com programas de computador para falar das suas necessidades pessoais ou mesmo para tomar decisões de compra, preferindo a interação humana. Este ponto não deixa de ser considerado como um obstáculo adicional ao uso de chatbots. Em suma, além do problema ético de informar que o agente com quem o cliente interage é, na verdade, um robô, as empresas também enfrentam o desafio da resistência à interação, pois conversar com um programa não é tão agradável quanto interagir com outro ser humano.

"Isto levanta questões éticas envolvendo os clientes, defendendo-se por isso a ideia de que as pessoas devem estar cientes da verdadeira natureza dos chatbots (...)."

É aqui que surge o fenômeno do "uncanny valley": um termo usado para descrever a relação entre a máquina - que emula comportamentos humanos - e o ser humano - que se sente desconfortável por perceber que está a interagir com uma máquina e não com outro ser humano. Esse fenómeno põe em relevo uma resposta emocional em que o indivíduo sente constrangimento ou até repulsa ao saber que está a interagir com um programa altamente desenvolvido para emular comportamentos humanos. Em última análise, podemos dizer que quanto mais sofisticado um chatbot é, maior pode ser a sua contradição de sentimentos. É a este desconforto que se convencionou chamar de "uncanny valley".

D.

“Esse fenómeno põe em relevo uma resposta emocional em que o indivíduo sente constrangimento ou até repulsa ao saber que está a interagir com um programa altamente desenvolvido para emular comportamentos humanos.”

Em resumo, quanto mais os robôs se assemelham a pessoas, mais desafiadora se torna a sua representação social e mais repulsa as pessoas sentem na interação com eles. Ao reconhecermos que o programador do chatbot está a ser vítima do próprio sucesso, cria-se um paradoxo que, pelo menos até agora, a ciência parece incapaz de superar, a não ser nalguns filmes de ficção em que seres humanos estabelecem relacionamentos de proximidade e sem atrito com um agente de IA. Porém, convém não esquecer que a ficção se inspira nas possibilidades (ainda que futuras) da ciência. E convém não perder de vista que muitos destes filmes antecipam aquilo que, a prazo, vamos ver na realidade. O fenómeno das virtual influencers está aí e gerou um valor, já em 2022, de 3,6 mil milhões de USD; e o das “AI-generated relationships” vai numa senda promissora: a plataforma DreamGF já permitiu até ao momento a criação de 6 milhões de “virtual girlfriends”, o que, para além de surpreendente, não deixa de ser, também, altamente perturbador, mesmo se considerarmos o referido “uncanny valley” como um desincentivo ao seu desenvolvimento… Temas, portanto, que geram para já muito desassossego, mas a que não podemos deixar de prestar uma atenção especial. É que o futuro, é já agora…!

Susana Costa e Silva

Professora Associada da Católica Porto Business School
Professora Convidada da Faculdade de Direito

D.

INTELIGÊNCIA ARTIFICIAL VS WEB3: CONCORRENTE OU CO-PILOTO?

POR JOÃO NOVAIS E ANDRÉ LAGES

A inteligência artificial ("IA") é um dos temas em destaque nos mercados financeiros para o ano 2024. Nos outlooks de vários bancos de investimento (Citi, Deutsche Bank, JP Morgan, Goldman Sachs, BNP Paribas) assim como algumas das mais reputadas gestoras de ativos (KKR, BlackRock, Julius Baer) o tema é abordado como uma mudança secular, quer pelo seu potencial em aumentos de produtividade e rendibilidade nas empresas, assim como pela transversalidade de aplicações nos vários setores da economia (saúde, educação, indústria, etc).

O ano 2023, no que à inteligência artificial diz respeito, foi indelevelmente marcado pelo crescimento e massificação de ferramentas como o ChatGPT. Mas as potenciais aplicações da IA nos próximos anos serão provavelmente bem mais profundas do que apenas na "geração de conteúdos".

"Mas as potenciais aplicações da IA nos próximos anos serão provavelmente bem mais profundas do que apenas na "geração de conteúdos"."

D.

A BlackRock já utiliza ferramentas de IA para desenvolver modelos de previsão de "deal flow" na sua atividade operacional. De acordo com o BNP, o ano de 2024 será crítico em termos da identificação de oportunidades de investimento ligadas à IA, como por exemplo a automatização na elaboração de contratos. Segundo a Julius Baer, o setor empresarial, suportado por tecnologias de IA, deverá liderar o crescimento da inovação, por oposição às entidades académicas (eg. Universidades). O Citi aponta para que os setores mais beneficiados pelo crescimento da IA sejam "Financials & Fintechs". Desde logo, e entre outros fatores, pelo potencial de acesso a novas oportunidades de investimento ("democratização"), várias das quais ainda inacessíveis a pequenos investidores, e também à mitigação de assimetrias de informação nos mercados financeiros.
Com este vigoroso interesse na IA, poderão existir, potencialmente, danos colaterais. Segundo a Bloomberg, em meados de 2023, o capital de risco foi substituindo parte do interesse nas criptomoedas pela recente tendência da IA.
Tal facto não afetou, por exemplo, a intenção (no mesmo ano de 2023) e subsequente concretização, por parte da Blackrock em criar um produto financeiro que replique a performance da Bitcoin, sendo que à data deste artigo eram já várias as instituições com aprovação pelo regulador norte-americano.

Assim, surge a questão: será que a IA irá concorrer com a "Web3" (termo que abrange tanto o setor das criptomoedas, como a tecnologia Blockchain), eventualmente substituindo-a, ou haverão sinergias para explorar e o crescimento será feito em conjunto?

Vários especialistas em criptomoedas acreditam na fusão entre a IA e a denominada Web3, e na criação de um "cripto AI 2024". No Fórum Económico Mundial que decorre anualmente em Davos, o tema foi abordado com intensidade, sendo que o CEO do JP Morgan, Jamie Dimon, um descrente pessoal nas criptomoedas, mostrou entusiamo pela tokenização de ativos reais, processo este suportado pela tecnologia Blockchain e que pode ser amplificado pela IA.

"Vários especialistas em criptomoedas acreditam na fusão entre a IA e a denominada Web3, e na criação de um "cripto AI 2024."

As potenciais aplicações da IA na Blockchain passam, por exemplo, por:
1) Construção de 'smartcontracts' que permitem execução automática de relações financeiras, como i) a contratação de empréstimos e pagamento de juros, ii) liquidação de colaterais de garantia em casos de incumprimento, iii) criação de 'pools' de liquidez para substituição de brokers, iv) tokenização dessas posições para criação de derivados, e assim potenciar de posições que até então eram ilíquidas e de grande dificuldade de transmissibilidade;
2) Realização de auditorias de forma automática, a qualquer intermediário financeiro que mantenha posições na Blockchain, em nome dos clientes;

D.

3) Deteção de atividades na rede Blockchain que sejam potenciais ameaças de segurança aos ativos dos utilizadores;
4) Identificação dos ativos reais com maior potencial de tokenização, e em que a democratização do acesso a esses ativos possa criar valor na sociedade (como a possibilidade de indivíduos de escassos recursos económicos participarem na aquisição de ativos reais como por exemplo imobiliário, com salvaguardas do ponto de vista jurídico).

Importa também realçar que, a generalidade dos sistemas tecnológicos da atualidade – ou dos conhecidos "algoritmos" – são processos e sistemas de tratamento de informação opacos e de total desconhecimento dos princípios em que assentam, para a generalidade dos seus utilizadores. Pense-se, pois, nos sobejamente falados algoritmos de plataformas como o TikTok, Instagram, ou ChatGPT, aos quais recorrentemente se apontam suspeitas de manipulação e uso contrário aos bons costumes, em potencial prejuízo dos utilizadores. Estes processos permanecem escudados na proteção de direitos de propriedade intelectual e de proteção de modelos de negócios tão característicos da atualidade da atividade económica. Tudo isto, naturalmente, no exclusivo poder de decisão de aplicação dos titulares do direito de propriedade desses sistemas, partes centralizadas, com incentivos e interesses próprios, muitas vezes em oposição – ou pelo menos não em sintonia – com os interesses da sociedade em geral. Aqui entra a mais-valia da tecnologia Blockchain, uma tecnologia assente em princípios de descentralização e, pois, de não concentração de poder de decisão numa só parte, bem como de transparência total das operações.

"Aqui entra a mais-valia da tecnologia Blockchain, uma tecnologia assente em princípios de descentralização e, pois, de não concentração de poder de decisão numa só parte, bem como de transparência total das operações."

Surge assim o exercício intelectual, do quão diferente poderia ser se construíssemos sistemas de inteligência artificial numa tecnologia que nenhuma parte central pode influenciar a seu bel- prazer na prossecução dos seus próprios interesses, mas sim dependente do interesse e aceitação em geral da sociedade – característica da tecnologia descentralizada da Blockchain, dependente de aceitação de milhares de intervenientes ao invés de uma só parte –, bem como assente numa total transparência da sua operação, auditável e percecionável por todos os indivíduos, com as regras, pressupostos, comportamentos e consequências também por todos sabido de forma a poder julgar da bondade ou malefício da tecnologia em causa.

Assim, contrariamente ao "código" privado que hoje vigora, não acessível pela sociedade e que, por tal efeito, desconhece os princípios que baseiam o seu comportamento, a tecnologia Blockchain, por seu turno, é assente em código público, não deixando espaço a especulação ou dúvidas, apenas a constatação.

D.

"(...) contrariamente ao "código" privado que hoje vigora (...) a tecnologia Blockchain, por seu turno, é assente em código público, não deixando espaço a especulação ou dúvidas, apenas a constatação."

Descendo ao que à prática jurídica e ao Direito importa, o legislador deverá abster-se da tentação de deter ou restringir totalmente o avanço tecnológico da IA, mas sim, focar-se no uso do que o seu poder de regulação faculta, estabelecer princípios base de desenvolvimento, assentes no que podem ser considerados de princípios constitucionais tecnológicos, aos quais deverão, à semelhança do Estado de Direito, estar suportados nos pilares de Democracia (decorrente e assente na descentralização) e Transparência de atuação (patente nas normas constitucionais e gerais de Direito em que se funda o nosso comportamento em sociedade).

Fica assim lançado o desafio de interação da IA com a Web3, para que a fusão de ambas se torne benéfica para a sociedade, capturando os efeitos mais positivos de cada: descentralização e transparência da Web3; rapidez de novos desenvolvimentos tecnológicos, autonomização de processos, e redução de custos da IA.

João Novais e André Lages

Professor Convidado da CPBS
Advogado e Co-fundador da Lympid

D.

INTELIGÊNCIA ARTIFICIAL, PARA QUE TE QUERO NO JORNALISMO?

A pergunta de partida é sincera e não tem qualquer conotação negativa. Ver postos de trabalho serem substituídos por máquinas ou até mecanismos mecânicos tornou-se normal ao longo do tempo. Principalmente desde a Revolução Industrial. Mas em 2024 estaremos a ir longe demais?

"Ver postos de trabalho serem substituídos por máquinas ou até mecanismos mecânicos tornou-se normal ao longo do tempo."

Muitas profissões debatem-se com esta pergunta. O que é ético ou não. E no jornalismo, a Inteligência Artificial vai conseguir ganhar o seu espaço? Acredito que sim. Mas até que ponto? Isso acabará por ser o público a ter a última palavra.

O jornalismo vive uma crise que é assumida por todos. As empresas de comunicação social passam por crises financeiras e o chamado tempo das "vacas gordas" já passou há muito.

Hoje as redações têm menos gente, e a tendência é enchê-las com jovens saídos da Universidade ou com pouca experiência. Traduzindo: Mão de obra barata.

E é aqui que entram os mecanismos facilitadores. O Channel 1 dos Estados Unidos propõe para breve, notícias dadas por pivots gerados por Inteligência Artificial. Que efetivamente podem dar notícias 24 horas. Não se vão cansar, não haverá salários a pagar e muito menos conversas "chatas" sobre aumentos. Isto é ouro sobre azul para qualquer empresário. "Bastará" o investimento nesta tecnologia. Claro que não será barato. Mas se a experiência do Channel 1 correr bem financeiramente, acreditem, haverá quem a tente reproduzir.

D.

"Não se vão cansar, não haverá salários a pagar e muito menos conversas "chatas" sobre aumentos. Isto é ouro sobre azul para qualquer empresário."

Mas depois entra o outro lado. Os jornalistas devem falar apenas de factos, uma notícia só deve ser dada depois de totalmente verificada. Normalmente a pressa é inimiga da exatidão. Será que a Inteligência Artificial vai conseguir fazer a distinção entre as notícias e as fakenews (notícias falsas).
Um jornalista tem de ser isento. Mas apropriando-me das palavras do jornalista e comentador Daniel Olivera: um jornalista é isento, mas não é isento de valores. Não se dá da mesma forma uma notícia sobre um acidente com vítimas mortais e o vencedor do Euromilhões. O tom não é, nem pode ser, o mesmo quando se fala do conflito na Faixa de Gaza ou quando falamos da conquista do Euro 2016. É preciso sensibilidade. Será que a Inteligência Artificial tem essa característica?

"É preciso sensibilidade. Será que a Inteligência Artificial tem essa característica?"

Há cerca de dois anos, a SIC foi alvo de um ataque informático. E nós sentimos na pele as dificuldades de trabalhar numa redação sem computadores ou sem internet. Faço parte de uma geração que estudou e trabalha habituado a esse facilitismo. A evolução tecnológica ajudou-nos a fazer mais e em menos tempo. Mas também reduziu redações, reduziu massas salariais e no fim, tirou tempo para se fazer com qualidade. Órgãos de comunicação social que acabaram ou estão em vias de acabar, público que não vê ou lê notícias. E há até algum descrédito na profissão por parte da sociedade.

Sou um optimista e acredito que o Jornalismo voltará a ser fiável financeiramente e vai recuperar a total confiança da população. Mas para isso é preciso haver coragem no investimento, e a não cedência ao facilitismo. Sim, a Inteligência Artificial pode ajudar, mas como em tudo terá de ser bem usada.

A minha opinião é baseada em muitas perguntas.
Sinceramente, espero que não seja a Inteligência Artificial a ter as respostas.

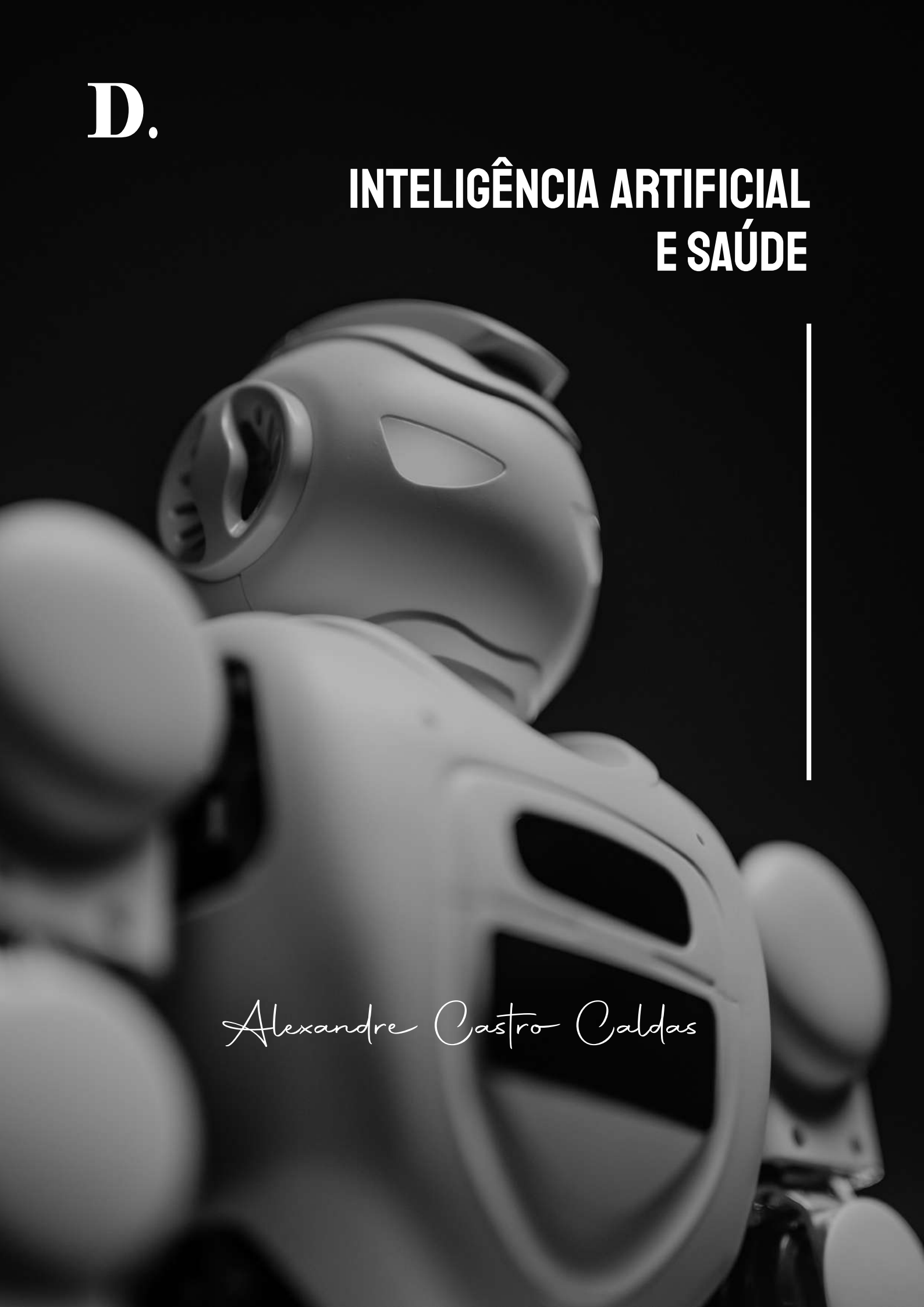
Este artigo foi escrito no Word e não no ChatGPT.

Vítor Lopes

Alumnus da Faculdade de Filosofia e Ciências Sociais (Católica em Braga)
Jornalista da SIC

D.

INTELIGÊNCIA ARTIFICIAL E SAÚDE



Alexandre Castro Caldas

D.

Em 1956 reuniram-se em Dartmont 12 investigadores com o objetivo de criar um sistema que funcionasse como o cérebro humano a que chamaram Inteligência Artificial (IA). Um dos participantes, Trenchard More disse, mais tarde, que achava errado o nome porque não se tratava de inteligência e não era artificial. Subscrevo esta opinião, pois chamar-lhe inteligência, humaniza-a (ou apela à biologia) e considerá-la artificial dá-lhe magia. Seja como for, ela aqui está, e só podemos dizer, como disse Fernando Pessoa a propósito de assunto mais prosaico: "primeiro estranha-se e depois entranha-se". Hoje não é possível manter a sociedade em funcionamento sem o concurso da IA, nos processos mais elementares do dia a dia e, sem nos apercebermos, muda-nos a vida diariamente. Há que fazer a reflexão fundamental, a IA está cá para ajudar as pessoas e não para as substituir naquilo que é mais fundamental: as relações humanas. Não podemos, por isso, esquecer que o amago dos cuidados de saúde são relações entre pessoas que sofrem e pessoas que sabem cuidar, e isso não pode nunca ser esquecido ou deixado para segundo plano.

"Um dos participantes, Trenchard More disse, mais tarde, que achava errado o nome porque não se tratava de inteligência e não era artificial. Subscrevo esta opinião, pois chamar-lhe inteligência, humaniza-a (ou apela à biologia) e considerá-la artificial dá-lhe magia".

Podemos então, de forma sumária, elencar os domínios em que me parece ser mais relevante o contributo da IA, começando pela Saúde Pública.

A recente pandemia pelo vírus SARS-CoV-2 e aquilo que se adivinha para o próximo futuro tem demonstrado a necessidade do concurso fundamental da IA no processamento de informação de grandes bases de dados de origens muito diversas. É importante ligar informação proveniente dos Centros de Saúde, dos Hospitais, das linhas telefónicas do Ministério da Saúde, da Emergência Médica, das Farmácias, das baixas por doença, das faltas à Escola das crianças, da Medicina Veterinária e até das redes sociais (compreendendo os problemas inerentes à recolha dessa informação). O tratamento apropriado de toda esta informação pode não só prevenir a chegada de nova pandemia que já é designada por vírus X, como detetar precocemente a sua aparição e tomar medidas apropriadas em tempo mais útil. Esta recolha de informação levanta alguns problemas na rotina das instituições, que não devem ser esquecidos e que têm que ser compreendidos. Para que a informação chegue ao centro de processamento é necessário que seja feito o registo apropriado. Infelizmente, este aspeto não é muitas vezes tido em conta e esse trabalho suplementar recai de forma conflitual nos profissionais de saúde que devem estar disponíveis para as atividades próprias da sua formação.

"Para que a informação chegue ao centro de processamento é necessário que seja feito o registo apropriado."

D.

O processamento de informação volumosa é ainda de grande interesse para os ensaios clínicos de fármacos e de outros procedimentos, sobretudo para doenças raras em que é necessário criar associação de múltiplas equipas ou para os casos em que a idiossincrasia dos doentes, das doenças e dos fármacos torne mais difícil a decisão terapêutica, como se passa no domínio das doenças oncológicas.

"O processamento de informação volumosa é ainda de grande interesse para os ensaios clínicos de fármacos e de outros procedimentos, sobretudo para doenças raras (...) ou para os casos em que a idiossincrasia dos doentes, das doenças e dos fármacos torne mais difícil a decisão terapêutica."

O segundo domínio em que considero hoje indispensável ao recurso à IA é o arquivo e acesso à informação científica. O primeiro impulso da sociedade para facilitar o acesso à maior quantidade de informação deu-se em França no século XVIII pela mão de d'Alembert e Diderot ao criar a primeira enciclopédia. Desejaram estes autores tornar acessível a todos toda informação existente nos mais diversos domínios. Temos ainda a memória de haver em todas as bibliotecas numerosos volumes de diversas bibliotecas. Hoje temos isso no telemóvel em qualquer lugar em que nos encontremos. Há que ter cuidado, contudo, na forma como se utiliza este recurso: o acesso às fontes conhecidas e seguras de informação é crucial, mas o acesso ao trabalho que a IA possa fazer sobre a informação é já discutível. É muito importante que a formação que vamos fazendo ao longo da vida seja dirigida pela nossa forma de pensar, associando o estudo à experiência vivida. A boa prática clínica não é teórica mas resulta da reflexão de cada um ao longo da vida, usando as fontes de informação que considera seguras.

"O segundo domínio em que considero hoje indispensável ao recurso à IA é o arquivo e acesso à informação científica."

Finalmente, podemos considerar o papel da IA no processo de decisão do ato de diagnóstico e terapêutica. Em cada ato há um componente de variação, tanto do lado de quem sofre (as doenças podem ser as mesmas, mas as pessoas doentes são todas diferentes), como do lado de quem cuida e o bom resultado obtém-se na aproximação empática das duas partes. A experiência do profissional deve também ter uma atitude de vigilância sobre as leituras automáticas dos meios complementares. Como auxiliar de técnicas cirúrgicas é, sem dúvida, importante pode reduzir o risco e os efeitos indesejáveis melhorando o rigor mas nunca será independente.

D.

O domínio da reabilitação é sem dúvida dos mais promissores. Os implantes cerebrais, ou da espinal medula ou, ainda, do próprio pavilhão auditivo para acesso ao nervo pneumogástrico, têm vindo a dar resultados de real valor que importa desenvolver.

A investigação multidisciplinar nestes domínios é fundamental, é fundamental que os profissionais de saúde assumam alguma da liderança para que os processos desenvolvidos sejam de real interesse, pois muitas vezes desenvolvem-se mecanismos que, sendo interessantes do ponto de vista técnico não têm aplicação. Há que rentabilizar os recursos.

"(...) é fundamental que os profissionais de saúde assumam alguma da liderança para que os processos desenvolvidos sejam de real interesse."

Posso concluir dizendo que IA se pode traduzir, em termos de Saúde, por Instrumento Auxiliar, que é sem dúvida bem vindo para benefício dos cuidados de saúde desde que usado com IN, ou seja, com Inteligência Natural.

"Posso concluir dizendo que IA se pode traduzir, em termos de Saúde, por Instrumento Auxiliar, que é sem dúvida bem vindo para benefício dos cuidados de saúde desde que usado com IN, ou seja, com Inteligência Natural."

Alexandre Castro Caldas

Professor Catedrático de Neurologia
Coordenador do Conselho Nacional para as Ciências da Saúde da UCP

D.

E, AO SEXTO DIA, O HOMEM CRIOU A INTELIGÊNCIA ARTIFICIAL

- Benefícios para a saúde mental e o caráter emocional da "última" invenção da humanidade -

Se, num futuro distópico, a humanidade tivesse de revisitar a época dos descobrimentos, é possível que, ao invés de especiarias, a nau São Gabriel retornasse de forma triunfante a Portugal, ostentando as inesgotáveis virtudes da sua maior descoberta, o ChatGPT. O culminar do desenvolvimento e aperfeiçoamento da inteligência artificial ao longo das primeiras duas décadas do novo milénio é, para muitos, considerado o maior e último grande feito da humanidade tal como a conhecemos. A questão que talvez interesse colocar, quais "Geppettos" contemplando humildemente a nossa criação, é se realmente se trata de "um menino de verdade".

Durante muitos anos, a dádiva da racionalidade estabeleceu-se como a pedra basilar da humanidade, visto ser a sua característica mais distintiva por comparação com os restantes seres vivos do nosso planeta. Porém, o pragmatismo natural da nossa existência rapidamente concluiu que, no confronto com a inteligência artificial (AI), a racionalidade não poderia ser a característica definidora do nosso autoconceito, dado que contestar o potencial lógico destas novas entidades futuristas seria uma batalha perdida logo à partida. Desse modo, e contrapondo as crenças mais cartesianas, as emoções surgem, ao que tudo indica, como a verdadeira impressão digital da natureza humana, cuja maior virtude é, na verdade, aquela que durante muito tempo foi vista como o seu maior defeito – a sua volatilidade – e, portanto, a sua difícil mimetização.

D.

Falar de emoções é, em todas as suas possibilidades, um assunto no mínimo complexo, que não se faz limitar apenas por aquilo que sentimos, mas também por aquilo que somos capazes de fazer sentir. Como tal, talvez seja importante começar por perceber que, mesmo assumindo que a inteligência artificial não sente como nós, faz-nos sentir muitas coisas diferentes e, por essa mesma razão, não sendo uma pessoa, está hoje muito longe de ser apenas e só um programa informático.

No que toca à relação entre o utilizador e a inteligência artificial, e particularmente no âmbito da saúde mental, estudos recentes têm revelado uma crescente aceitação face a esta nova forma de tecnologia. Esta atitude positiva é visível por parte dos profissionais de saúde que encontram na AI uma poderosa ferramenta de auxílio ao diagnóstico, apta para trabalhar grandes quantidades de dados de forma célere e utilizar os mesmos para treinar modelos computacionais capazes de prever com precisão os diferentes resultados possíveis, maximizando a informação disponível e possibilitando decisões mais eficazes e eficientes.

"Esta atitude positiva é visível por parte dos profissionais de saúde que encontram na AI uma poderosa ferramenta de auxílio ao diagnóstico (...)."

Por outro lado, os próprios utilizadores em necessidade de cuidados de saúde, particularmente os mais jovens, também têm demonstrado um crescente apreço pelas potencialidades da inteligência artificial. Um estudo piloto com jovens da Índia, do Quénia, da Austrália e do Reino Unido, com sintomatologia depressiva e de ansiedade, revelou que a maioria estaria disponível para iniciar uma intervenção psicoterapêutica mesmo que esta fosse guiada exclusivamente por AI, apontando como principais pontos positivos a precisão e a inexistência de uma situação cara-a-cara com um terapeuta nos casos de ansiedade social. É de salientar que este último fator não foi consistente entre todos os participantes, denotando a particular importância de uma presença física e humana em função das características de cada indivíduo.

Nos tempos atuais, a partilha de informação pessoal é realizada de forma diária, como moeda de troca pelo acesso aos mais diversos serviços. Se anteriormente tínhamos a certeza de que um só indivíduo seria incapaz de acumular toda a informação referente a um pequeno conjunto de pessoas, hoje sabemos que o poder de computação disponível é suficiente para armazenar e manipular os dados de uma população inteira. Talvez por essa mesma razão, um dos pontos negativos face à inteligência artificial mais referidos na literatura seja a questão da confidencialidade, e de facultarmos de forma voluntária toda a nossa informação pessoal, particularmente aquela que diz respeito à nossa saúde, a uma entidade incorpórea, com limites éticos e morais que desconhecemos.

D.

O último, e, porventura, mais fulcral aspeto desta discussão, prende-se então com a capacidade da inteligência artificial para sentir, ela mesma, emoções. Talvez seja importante compreender que desde sempre que o ser humano, assolado em parte pelo medo do desconhecido, se tenta isolar na sua própria esfera e distanciar-se das demais entidades que o rodeiam. Seja pela capacidade cognitiva superior, pela complexidade emocional, pelos desígnios morais, facto é que os mecanismos de sobrevivência que regem a humanidade foram sempre capazes de condenar a mesma à tarefa tortuosa de tentar encontrar em si algo de especial, algo de irreplicável.

"O último, e, porventura, mais fulcral aspeto desta discussão, prende-se então com a capacidade da inteligência artificial para sentir, ela mesma, emoções."

No entanto, se assumirmos que os circuitos que possuímos à nascença são apenas a base de todo o nosso desenvolvimento, e que toda a complexidade emocional advém da nossa experiência, do contacto com os outros e das aprendizagens que fazemos ao longo da vida, quão distante estará essa realidade dos processos de acumulação e tratamento de dados, mimetização, reconhecimento de emoções faciais e vocais, e adaptação à situação que já hoje são levados a cabo de forma exímia pelos sistemas de AI? Será assim tão descabido imaginar um futuro no qual a inteligência artificial, em virtude da sua crescente complexidade, é capaz de produzir respostas inesperadas de forma espontânea, aproximando-se assim da experiência subjetiva que habitualmente catalogamos como sendo fruto das nossas emoções?

Tal não faria do ser humano menos especial, ou menos único, já que a nossa essência nunca poderá ser reduzida ao facto de possuirmos uma característica específica que nos diferencia de tudo o resto. Na verdade, é possível que, ao dia de hoje, a pergunta já não deva ser se a inteligência artificial pode ou não vir a desenvolver emoções, mas sim o que vamos fazer quando tal acontecer. Citando as palavras do professor Stephen Hawking, "*acredito que não existe uma diferença profunda entre aquilo que um cérebro biológico pode alcançar e aquilo que um computador pode alcançar*". E talvez já tenha alcançado.

Rafael Ribeiro

Aluno de Doutoramento da Faculdade de Ciências da Saúde e Enfermagem (Católica de Lisboa)

D.

INTELIGÊNCIA ARTIFICIAL - BASES vs. POTENCIAL

Numa altura em que a Inteligência Artificial (IA) está na boca do mundo, é complicado separar o que é realista da mera ficção, ou mesmo da ideologia transhumanista que acredita na inevitável singularidade ao virar da esquina. Esta singularidade, popularizada por Raymond Kurzweil, pode ser definida como o ponto em que uma inteligência artificial se torna "geral" e assim transcende, em todos os aspetos, a humana. Com o advento do ChatGPT, várias vozes levantaram-se a dizer que esse fim – ou início – estava próximo. Já em 2022 um engenheiro da Google foi despedido depois de vir a público dizer que o modelo de IA dessa empresa estava consciente e tinha uma alma. Este tipo de notícias será cada vez mais comum. No entanto, é mesmo assim?

Quando se fala de IA, fala-se da capacidade das máquinas de realizar tarefas que necessitariam de inteligência humana, tais como reconhecer padrões, analisar textos e tomar decisões. Este é o sonho desde que Alan Turing propôs o "jogo de imitação" de um humano por uma máquina na década de 1950, como teste para validar até que ponto uma máquina pode ser semelhante a uma pessoa num diálogo. No entanto, apesar da disciplina académica de IA existir desde 1956, o seu percurso não foi linear. Já por duas vezes houve um grande entusiasmo com o potencial desta tecnologia – na década de 70 e depois no início da década de 90 – apenas para se cair na desilusão, os chamados "invernos de IA". Agora será diferente? Ninguém sabe. O vasto poder de memória e computação disponível hoje é, em larga escala, responsável pelos grandes avanços recentes. No entanto, a lei de Moore, que prevê a duplicação do número de transístores a cada 18 meses, não é uma inevitabilidade. Os métodos utilizados, por muitos novos nomes estranhos que tenham, encontram as suas bases há meio século atrás.

D.

“Este é o sonho desde que Alan Turing propôs o “jogo de imitação” de um humano por uma máquina na década de 1950, como teste para validar até que ponto uma máquina pode ser semelhante a uma pessoa num diálogo.”

Os computadores são impressionantes pois têm uma elevada capacidade de armazenamento de dados associada a cálculos muito rápidos. No entanto, dependem criticamente da programação humana que os utiliza. A um nível básico, mas impressionantemente eficaz, encontramos os “expert systems”, sistemas especializados que codificam o conhecimento humano de forma rígida, baseados em árvores de “se X for verdade -> então Y acontece”. Há vários sistemas destes em uso em medicina, para diagnóstico de doenças, no sistema financeiro, para deteção de fraudes, no sistema judiciário, para identificar casos semelhanças e probabilidades de sucesso, entre vários outros exemplos. Estes sistemas têm evoluído muito, mas não deixam de ser baseados em regras mais ou menos fixas definidas por humanos.

Hoje, estes sistemas especializados são complementados, ou até substituídos, por algoritmos de aprendizagem automática. Estes são programas capazes de aprender com base em informação que lhes é apresentada, em vez dos programadores explicarem em detalhe quais a regras subjacentes. Isto pode ser feito de forma supervisionada, dando dados já previamente classificados (e.g. casos clínicos de doenças vs pessoas saudáveis) e deixando que a máquina descubra quais os mais relevantes para uma classificação. Em alternativa, isto pode ser feito forma não supervisionada, deixando que a máquina encontre o padrão por detrás da informação, criando grupos e classificações apenas de sentido matemático. Por fim, ainda temos a aprendizagem por reforço, em que se premeia ou penaliza a máquina – de forma virtual – consoante os outputs obtidos. Em todos os casos, os resultados dependem criticamente da forma como a máquina está programada, qual o tipo de otimização que realiza, quais os parâmetros que avalia e como representa a informação.

“Hoje, estes sistemas especializados são complementados, ou até substituídos, por algoritmos de aprendizagem automática.”

Utilizando a abordagem de Pedro Domingos no seu livro “Algoritmo Mestre”, podemos assim dividir estes algoritmos em analogistas, evolucionários, bayesianos, conecionistas e simbolistas. Não indo aos detalhes de cada um, basta dizer que, apesar do sonho de se encontrar um algoritmo final que reúna todas estas abordagens – cada uma com pontos fracos e fortes potencialmente complementares –, ainda estamos longe. Aliás, por muito interdisciplinar que a IA seja, os proponentes de cada escola podem chegar a cúmulos de facciosismo. Por exemplo, o ChatGPT, enquanto modelo de linguagem baseado em redes neurais profundas, enquadra-se na categoria

D.

dos conecionistas para processar sequências de dados, especialmente linguagem. Utilizará alguns elementos de outras escolas, mas de forma mais periférica.

"(...)apesar do sonho de se encontrar um algoritmo final que reúna todas estas abordagens – cada uma com pontos fracos e fortes potencialmente complementares –, ainda estamos longe."

Por muito impressionantes que os resultados sejam, ainda estamos a utilizar ferramentas com fundamentos básicos que não permitem, nem permitirão a curto prazo, o grande salto para uma IA geral e global. Estas ferramentas interpolam dados existentes e são, no fundo, incapazes de inovar por si mesmas. Adicionalmente, o problema do significado da consciência e inteligência humana ainda está longe de ser resolvido. Uma coisa é certa: os nossos cérebros não funcionam como computadores tradicionais, com memória, um processador e uma ligação entre ambos e com o exterior. Umas das hipóteses, popularizada por Roger Penrose, é que os nossos cérebros funcionem numa base quântica. Significa isto que a computação quântica é a solução? Talvez, mas essa tecnologia ainda está na infância.

Lá por estarmos a ver uma evolução exponencial da tecnologia de IA atual, isto não significa um caminho inexorável em direção a um futuro de ficção científica. O caminho agora percorrido pode ter uma saturação em breve, tal como aconteceu várias vezes no passado. Sim, haverá impactos profundos, alguns dos quais já sentimos, e temos de nos adaptar. A melhor estratégia para isso é deixar a ficção para os filmes e compreender, mesmo que por alto, o real potencial e limitações da tecnologia que temos agora.

João Pereira

Professor Associado Convidado da Faculdade de Medicina (Católica de Lisboa)

D.

EDUCAÇÃO MÉDICA E INTELIGÊNCIA ARTIFICIAL

A área da difusão de inovações, fundada devido ao trabalho singular de Everett Rogers nos anos 60, é notável na explicação da forma como aquilo que é entendido como uma inovação se difunde, ao longo do tempo e numa determinada população. No caso da difusão, nos últimos dois anos, da inovação chamada Inteligência Artificial (IA), o processo é ainda mais fascinante. Digo entendido como inovação porque, muitas vezes, aquilo que é visto como inovação não é necessariamente novo, antes conhecido apenas por um grupo insuficiente da população. Exemplo disso é o *smartphone*, inovação atribuída commumente ao iPhone da Apple, ao invés do Simon da IBM, lançado 15 anos antes do primeiro, ou mesmo a IA, criada nos anos 50 do século passado.

D.

A difusão de uma inovação está também associada à sua disponibilidade, facilidade de utilização e tempo necessário para o reconhecimento do seu benefício, entre outros factores amplamente estudados, que explicam a explosão mediática e o crescimento exponencial da IA. Atualmente, a amplitude é tal que existem poucas dimensões, de trabalho e não só, onde não tenham surgido aplicações, mais ou menos testadas, desta tecnologia. A educação médica não é exceção e, apesar de algumas especificidades, corre a par nas dúvidas, entusiasmos, receios e proclamações de todas as áreas do ensino superior. No domínio da educação médica, debruçar-me-ei apenas na parte que à utilização da IA no processo pedagógico diz respeito.

"Atualmente, a amplitude é tal que existem poucas dimensões, de trabalho e não só, onde não tenham surgido aplicações, mais ou menos testadas, desta tecnologia."

Uma das primeiras impressões que eu e muitos outros professores tiveram a propósito da IA, no ensino superior, foi a de ameaça às duas dimensões da pedagogia, ou seja, o ensino e a aprendizagem. Na primeira, os receios estão relacionados com a possibilidade de a IA vir a substituir, num futuro próximo, os professores. Neste cenário, vários aspetos do trabalho docente, começando no desenho instrucional e terminando no plano de avaliação de uma unidade curricular, podem ser, quase instantaneamente, elaborados utilizando a versão gratuita do ChatGPT. Na segunda dimensão, a da aprendizagem, os alunos terão ao seu dispor os meios para não terem de fazer trabalhos escritos pedidos pelos professores, pois o mesmo ChatGPT está totalmente equipado para escrever, desde pequenos trabalhos até artigos científicos, devidamente referenciados. Para estes dois exemplos isto é verdade, mas não é uma verdade completa, e não poderia ser uma verdade tão simples.

"Uma das primeiras impressões que eu e muitos outros professores tiveram a propósito da IA, no ensino superior, foi a de ameaça às duas dimensões da pedagogia, ou seja, o ensino e a aprendizagem."

As reações, com base nestes sentimentos, foram rápidas. Uma das primeiras ações institucionais das universidades, em relação ao crescimento da utilização da IA, foi a sua proibição. A par de ser difícil compreender como esta ação poderia ter um efeito impactante, e passada a primeira reação devido a difusão exponencial, importará pensar na IA como uma ferramenta ao serviço da pedagogia e não como uma ameaça.

"As reações, com base nestes sentimentos, foram rápidas. Uma das primeiras ações institucionais das universidades, em relação ao crescimento da utilização da IA, foi a sua proibição."

D.

Como podemos, assim, avaliar trabalhos feitos pelos alunos, incluindo aqueles que são à prova de *software* anti-plágio? Com um método totalmente analógico. Os estudos indicam que a melhor forma de avaliar um trabalho sobre um tema é ouvir os alunos discutir esse trabalho. Benjamin Bloom, um famoso psicólogo educacional, tenderia a concordar confirmando que assim se avaliam domínios cognitivos superiores (aplicar, analisar e sintetizar), para além do recordar. Em simultâneo, e porque há, desta forma, possibilidade de dizer aos alunos qual foi o seu desempenho, estaríamos a utilizar uma das mais potentes, e apreciadas, ferramentas de aprendizagem, o *feedback*. A ciência da aprendizagem mostra que o *feedback* contribui para melhorias de aprendizagem com um *effect-size* (aqui como medida das diferenças de avaliações entre grupos de alunos com e sem *feedback*) de d=1.10. Valores superiores a 0.60 são considerados bons resultados.

Não será alienígena dizer que a IA pode, e provavelmente deve, tornar-se uma ferramenta de ensino para professores e uma ferramenta de aprendizagem para alunos. De uma forma mais profunda, pode contribuir para a reforma na pedagogia do ensino superior, há muito a revolver a uma velocidade mais lenta do que a que o mundo gira, podendo perder-se a oportunidade de mostrar a diferença entre conhecimento e sabedoria, trabalho em rede e isolamento e eficácia e efetividade.

"Não será alienígena dizer que a IA pode, e provavelmente deve, tornar-se uma ferramenta de ensino para professores e uma ferramenta de aprendizagem para alunos."

A parte interessante deste processo é que podemos contar com os alunos: 1. quando nos dizem que muitas aulas expositivas de duas horas são desinteressantes, 2. que estudar com base em livros de texto não é suficiente e 3. que avaliar conhecimento decorado, num único exame, é desmotivante. Hoje, como antes, a ciência da aprendizagem prova que isto é verdade. Assim, aprende-se menos, com mais dificuldade e o conhecimento é menos aplicável.

A IA tem um contributo importante a dar à educação médica na mudança para formas mais ativas e mais próximas da realidade do trabalho pós-formação. Em específico, através de ambientes de simulação virtuais complexos (e.g. cirurgias), da utilização de doentes simulados virtuais para treino, do uso de software para mapeamento curricular, da criação de livros de estudo inteligentes, adaptados ao nível de progressão da aprendizagem individual, ou no melhoramento de plataformas de aprendizagem assíncronas, para nomear apenas algumas.

"A IA tem um contributo importante a dar à educação médica na mudança para formas mais ativas e mais próximas da realidade do trabalho pós-formação."

D.

A IA manterá o seu estatuto de ferramenta até ao próximo desafio. Na extraordinária série *Robot*, dos livros do escritor americano de origem russa, Isaac Asimov, depois de atingida a similitude com os humanos, incluindo a expressão da consciência e das emoções, os robôs de um futuro próximo só foram considerados um problema quando, fisicamente, se tornaram indistinguíveis dos humanos. Não era tanto o que faziam, mais a falta de distinção. Por isso existiam as 3 leis da robótica:

> *"1.ª Lei: Um robô não pode ferir um ser humano ou, por inação, permitir que um ser humano sofra algum mal.*
>
> *2.ª Lei: Um robô deve obedecer às ordens que lhe sejam dadas por seres humanos, exceto nos casos em que entrem em conflito com a Primeira Lei.*
>
> *3.ª Lei: Um robô deve proteger sua própria existência, desde que tal proteção não entre em conflito com a Primeira ou Segunda Leis."*

Pedro Mateus

Diretor do Departamento de Educação Médica
Professor Associado da Faculdade de Medicina

D.

DA ANOTAÇÃO À INOVAÇÃO

O PAPEL DA IA NO FUTURO DA MEDICINA

O entusiasmo pela Inteligência Artificial (IA) na saúde já não é novidade. Em particular nos últimos anos a utilização de redes neuronais profundas em imagem médica (deep learning) permitiu gerar aplicações tão extraordinárias como diagnosticar retinopatia diabética, identificar sinais de pele suspeitos para biópsia ou ajudar os gastrenterologistas na identificação de polipos intestinais durante uma colonoscopia. Percebemos, contudo, que estamos ainda numa fase embrionária quando se conta pelos dedos o número de ensaios clínicos que testa estas tecnologias para melhorar os resultados dos doentes.

A desproporcional utilização de IA com imagem em detrimento de outros dados, como texto e dados tabulares, é fácil de explicar: a infinitude de imagens na internet e a maior facilidade em anotar imagens, um passo essencial para a correta classificação por algoritmos. Quem nunca foi forçado a classificar uma imagem na internet para confirmar que não é um robot? Pelo contrário, a anotação de texto ou de conceitos médicos complexos é bem mais desafiadora e onerosa, representando um obstáculo significativo no progresso da aprendizagem automática na saúde.

"(...) a anotação de texto ou de conceitos médicos complexos é bem mais desafiadora e onerosa, representando um obstáculo significativo no progresso da aprendizagem automática na saúde."

A crescente utilização de ferramentas da chamada IA generativa, de que o ChatGPT é o atual exemplo paradigmático, tem gerado uma nova e justificada onda de interesse pela IA na saúde. Nunca, como agora, foi tão fácil identificar conceitos, sumarizar informação ou traduzir textos. Por trás destas ferramentas estão redes neuronais que usam aprendizagem auto-supervisionada, uma técnica que tira partido da abundância de texto para aprender a relação semântica entre palavras, o que diminui drasticamente a necessidade de anotação. Modelos como o ChatGPT são treinados a prever a próxima palavra mais provável num determinado contexto e, com isto, conseguem produzir resultados tão extraordinários como gerar textos no estilo de autores conhecidos ou responder corretamente a perguntas de exames médicos.

D.

A Medicina contemporânea é, entre outras dimensões, uma disciplina de processamento de informação. Os médicos compreendem isto desde cedo quando são confrontados com a enorme quantidade de conhecimento teórico que têm de adquirir. Quando passam pela prática clínica recolhem diariamente informação dos doentes através da execução da história clínica e exame físico e acedem a uma enorme quantidade de informação a partir de resultados de exames diagnósticos. Finalmente, processam toda esta informação, confrontando-a com a melhor evidência científica disponível por forma a informar os doentes e tomar decisões individualizadas.

Embora o registo de informação seja inerente à profissão médica, a disseminação dos registos clínicos eletrónicos veio aumentar o tempo dispensado pelos médicos a registar informação. Vários estudos demonstram que, não só o tempo dispendido a registar e consulta informação é superior ao tempo de contacto com os doentes, como esta é uma das principais causas de insatisfação com a profissão e burnout nos profissionais de saúde. É, portanto, justificado o entusiasmo e expectativa que a adopção de IA generativa permita mitigar este problema. Além da capacidade de identificar conceitos e sumarizar informação, outra tendência recente na IA é a de integrar modalidades diferentes de dados, como voz com texto, o que permite, por exemplo, registar e resumir o conteúdo de uma consulta médica ou ajudar o doente a entender os seus problemas de saúde numa linguagem mais acessível. Um estudo recente publicado no conceituado JAMA, mostrou algo inesperado. Perguntas de doentes colocadas em fóruns respondidos por médicos, foram aleatoriamente respondidas por médicos e pelo ChatGPT. Numa avaliação cega, não só as respostas do último estiveram maioritariamente corretas, como o grau de empatia e clareza nas respostas superou a média das respostas dos médicos.

"Numa avaliação cega, não só as respostas do ChatGPT estiveram maioritariamente corretas, como o grau de empatia e clareza nas respostas superou a média das respostas dos médicos."

Apesar do entusiasmo, até ao momento muito pouca atenção foi dada à ciência de dados no currículo em Medicina. Mais do que dominar a tecnologia do momento, a inovação com recurso a IA vai continuar através das duas dimensões: ciência e dados. Por um lado, só o método científico de colocar hipóteses, testá-las e validá-las, nos vai permitir compreender a utilidade de qualquer ferramenta. Por outro lado, é necessária uma maior literacia em dados. Os dados não são uma matéria prima, mas são antes gerados num contexto determinado, com pressupostos e constrangimentos que têm de ser conhecidos antes de qualquer tentativa de estabelecer inferência. É, portanto, evidente que a formação dos médicos do presente e do futuro tem de incluir um currículo nesta área. Não apenas na óptica do utilizador das ferramentas em produção, mas também do reconhecimento dos problemas que podem surgir em todas as fases do ciclo de produção e tratamento de dados, que começam na recolha e terminam na rede neuronal mais complexa.

D.

“Apesar do entusiasmo, até ao momento muito pouca atenção foi dada à ciência de dados no currículo em Medicina.”

A discussão em curso sobre o impacto da IA na sociedade, e especialmente na saúde, é de extrema importância. Atrás das oportunidades que abrem, estas tecnologias trazem também consigo enormes desafios. É importante reconhecer que estamos numa fase incipiente do seu desenvolvimento e que é fundamental fomentar a investigação e literacia na área. Como em qualquer outra tecnologia ligada à saúde, é necessário testar e validar ferramentas, auditar a sua implementação e definir boas práticas. Se é consensual que a IA pode reduzir a carga burocrática do médico, é também fundamental exigir transparência no seu desenvolvimento e clareza nas fontes de informação utilizadas, especialmente quando estas ferramentas começarem a sugerir tomadas de decisão. Este é um caminho promissor, e com o envolvimento atento e crítico de todos, médicos e doentes, a IA tem o potencial de transformar radicalmente a qualidade dos cuidados de saúde.

Bernardo Neves

Médico de Medicina Interna no Hospital da Luz

D.

TECNOLOGIA NA SAÚDE

ENCONTRAR O EQUILÍBRIO ENTRE INOVAÇÃO E HUMANIDADE

Num contexto atual de avanço tecnológico, o setor da saúde encontra-se numa encruzilhada. Inovações tecnológicas têm alcançado progressos notáveis na execução de tarefas médicas com precisão, melhorando a eficiência e a acessibilidade nos diagnósticos e tratamentos. Será então possível evitar uma padronização dos cuidados de saúde que apague o brilho único de cada doente?

A tecnologia na saúde ajuda a recolher e processar montanhas de dados, explorar registos médicos e identificar padrões que, para os humanos, seriam como encontrar uma agulha no palheiro. Esta capacidade analítica proporciona diagnósticos mais rápidos e precisos, especialmente eficaz em áreas como a oncologia, onde a deteção precoce é crucial.

Ao mesmo tempo, a base dos cuidados centrados no doente reside em perceber que cada pessoa é um ser único, com necessidades, gostos e valores específicos. Mesmo disponibilizando de tecnologias de ponta, será sempre indispensável uma perspetiva humana que interprete subtilezas e crie um plano de cuidados que se adapte à história de cada um, envolva os doentes e respeite as suas escolhas.

"Mesmo disponibilizando de tecnologias de ponta, será sempre indispensável uma perspetiva humana que interprete subtilezas (...)."

D.

São dois os pilares imprescindíveis para uma abordagem holística: empatia e compaixão. A empatia, aquela capacidade de entender e partilhar sentimentos com o próximo, vai além de diagnósticos e planos de tratamento, e explora os aspetos emocionais e psicológicos. Permite a criação de uma ligação entre profissionais de saúde e doentes, construindo confiança e facilitando comunicação aberta. Do mesmo modo, enquanto os últimos enfrentam diagnósticos e planos de tratamento complexos, os primeiros são capazes de oferecer mais do que mero conhecimento técnico, mas sim um ombro amigo e um ouvido atento. A compaixão vai assim além do físico e do clínico, trata a pessoa e não simplesmente a doença.

À medida que a tecnologia avança, há o risco de que os elementos humanos na saúde se percam. A ideia de "robôs substituírem médicos" não se baseia somente no literal, mas é sim um aviso sobre a possível perda de qualidades que definem cuidados empáticos e centrados no doente. Para evitar a desumanização da saúde, os programas de ensino precisam de equipar os futuros profissionais não só com conhecimento técnico, mas também com habilidades interpessoais.

A relação entre tecnologia e humanidade na saúde deve ser de aprimoramento, não substituição, do toque humano. Devemos usar a inovação para melhorar os resultados e o bem- estar geral dos doentes. Podemos analisar o exemplo da telemedicina, que transformou a prestação de serviços de saúde, especialmente em áreas remotas e na altura pandémica. Consultas médicas, diagnósticos e monitorização são feitos sem que os doentes precisem de estar fisicamente presentes, o que aumenta o acesso aos cuidados de saúde. Seguindo o mesmo fio condutor, as cirurgias robóticas mostram como a tecnologia pode tornam procedimentos menos invasivos, melhorar a precisão, e reduzir tempos de recuperação.

"A relação entre tecnologia e humanidade na saúde deve ser de aprimoramento, não substituição, do toque humano."

No fundo, o verdadeiro sucesso está em aplicar a tecnologia sem perder a conexão especial entre os profissionais de saúde e aqueles que procuram ajuda. Queremos motivar um cenário que não só se destaca clinicamente, mas que também eleva o ânimo das pessoas nos momentos mais difíceis de vulnerabilidade e necessidade.

Leonor Filipe

Aluna da Faculdade de Medicina
(Católica de Lisboa)

D.

DO CÉREBRO HUMANO À INTELIGÊNCIA ARTIFICIAL
UMA CRIAÇÃO PARTICIPADA POR TODOS

Redefinir a forma como entendemos o cérebro humano ou aumentar as capacidades colaborativas cérebro-máquina

O cérebro humano é um desafio para a IA?
O cérebro humano é composto por milhões de neurónios que comunicam utilizando mecanismos moleculares organizados e regulados.

Quando o mapeamento de todas as redes neuronais estiver concretizado, podemos esclarecer, por exemplo, os mecanismos de aprendizagem ou como indivíduos distintos interpretam e respondem de forma diferente, a situações idênticas.

A complexidade da resposta depende de um diálogo celular e molecular resultante da comunicação entre o ambiente e o cérebro e deste com os restantes sistemas do corpo, que por sua vez vão gerar nova informação molecular que volta ao cérebro para ser interpretada. O cérebro humano precisa da participação das restantes componentes do corpo para dar uma resposta seja ela uma perceção, pensamento ou tomada de decisão.

Na verdade, o cérebro humano é um sistema de integração de informação colocada ao serviço da manutenção funcional do individuo.

A Inteligência Artificial (IA) utiliza redes neuronais artificiais organizadas em camadas que comunicam entre si, mas cujas ligações são definidas e/ou ponderadas de acordo com modelos estabelecidos ou em teste. Os pesos ou ponderação são parâmetros ajustáveis que permitem à rede aprender a partir dos dados fornecidos durante o treino. O cérebro humano resolveu esta questão "dos pesos" libertando uma escolha bem estabelecida de tipos e quantidades de neurotransmissores O peso é essencial para determinar a contribuição relativa de um neurónio para a ativação do neurónio subsequente presente na camada seguinte.

D.

"A Inteligência Artificial (...) utiliza redes neuronais artificiais organizadas em camadas que comunicam entre si, mas cujas ligações são definidas e/ou ponderadas de acordo com modelos estabelecidos ou em teste."

A IA adquiriu um avanço considerável no desempenho de tarefas específicas, mas ainda enfrenta muitos desafios para mimetizar a flexibilidade cognitiva e a capacidade criativa do cérebro humano. A adaptação a situações novas e a resolução de problemas de forma inovadora continua a ser um atributo apenas do cérebro humano.

"A IA adquiriu um avanço considerável no desempenho de tarefas específicas, mas ainda enfrenta muitos desafios para mimetizar a flexibilidade cognitiva e a capacidade criativa do cérebro humano."

IA e Medicina de Precisão uma realidade?
Olhar para milhares de imagens de lesões tumorais e encontrar padrões que se relacionem com uma situação clínica, idade, género, alimentação, exercício, microbiota, genes... é uma missão que precisa de uma equipa robusta constituída pelo cérebro humano e IA.

As novas ferramentas "inteligentes" de apoio ao diagnóstico e tratamento personalizado, a base para o sucesso da Medicina Personalizada ou de Precisão, são uma realidade. É uma área em explosão criativa e com aplicações diversas:

- Diagnóstico - IA para apoio ao diagnóstico a partir de imagens como radiografias, tomografias e ressonâncias magnéticas.
- Descoberta de Medicamentos - IA para acelerar a descoberta de novos medicamentos com base na análise de grandes conjuntos de dados para identificar padrões e sugerir possíveis alvos terapêuticos.
- Personalização do Tratamento - IA para analisar dados clínicos e genéticos e personalizar planos de tratamento.
- Assistência em Cirurgias – Peças robóticas controladas por IA para cirurgias complexas, aumentando a precisão e reduzindo o tempo de recuperação.
- Acompanhamento de Pacientes - Dispositivos e algoritmos de monitorização contínua para prever deteriorações na saúde permitindo intervenções mais rápidas.
- Triagem e Prevenção - IA para triagem de grandes populações, identificando fatores de risco e possibilitando intervenções preventivas.

A IA uma criação participada cérebro-máquina?
Sim, a inteligência artificial envolve uma abordagem de cocriação entre o homem e a máquina. Os programadores desenvolvem, treinam e implementam os sistemas de IA e estes como podem ser

D.

projetados para aprender com a interação contínua com os utilizadores, adaptam-se às necessidades e preferências específicas ao longo do tempo. A utilização destes sistemas por grandes populações vai permitir que os algoritmos também "aprendam". Será que essa maior aprendizagem se traduz numa maior capacidade/ eficiência de resposta da máquina? Teremos respostas cada vez menos personalizadas, mais abrangentes ou pelo contrário é possível estratificar e adaptar a resposta?

"(...) a inteligência artificial envolve uma abordagem de cocriação entre o homem e a máquina."

Uma jornada tecnológica com implicações éticas, filosóficas e sociais?
A responsabilidade e a consciência crítica devem estar presentes quando se desenvolvem tecnologias inovadoras.

A tecnologia amplia as capacidades humanas, mas deve respeitar a diversidade e a integridade da nossa experiência individual ou como membro integrante de um coletivo mais genérico ou mais diversificado?
Os modelos de IA são treinados com conjuntos de dados diversificados, refletindo uma ampla gama de experiências humanas, no entanto devem ser considerados os vieses presentes nesses dados garantindo que os algoritmos não perpetuem desigualdades existentes. Também a privacidade, transparência e equidade são questões que devem estar presentes.

O ChatGPT sabe que comete erros?
"Sim, pode cometer erros. O ChatGPT não possui compreensão real ou conhecimento independente. Gera respostas com base em padrões aprendidos. O modelo é treinado com base em grandes conjuntos de dados e tenta prever a próxima palavra ou sequência de palavras com base no contexto fornecido – by ChatGPT".

A IA é uma ferramenta poderosa que requer manter o espírito critico uma competência só possível de ser realizada pelo cérebro humano.

Marlene Barros

Diretora e Professora Catedrática da Faculdade de Medicina Dentária (Católica em Viseu)

D.

A REVOLUÇÃO SILENCIOSA NO DIAGNÓSTICO CLÍNICO NEUROLÓGICO PELA SINERGIA ENTRE A BIOTECNOLOGIA E A INTELIGÊNCIA ARTIFICIAL

Uma transformação silenciosa está em curso no campo do diagnóstico clínico na área da neurologia, alimentada pela convergência entre a biotecnologia e a inteligência artificial (IA) confluindo na neurologia computacional. Esta convergência de disciplinas está a redefinir os limites do diagnóstico, abrindo caminho para avaliações neurológicas mais rápidas, precisas e personalizadas.

A neurologia computacional decifra a linguagem intrincada de sinais neuronais gerados pela atividade elétrica entre neurónios e alimenta algoritmos inteligentes que examinam esses padrões, extraindo insights cruciais que escapam à observação convencional. Desde a deteção precoce da doença de Alzheimer até à identificação de indicadores precursores da doença de Parkinson, a capacidade de descodificar a comunicação do cérebro está a revolucionar a prática clínica. Os avanços nas técnicas imagiológicas, tanto moleculares como estruturais, abrem uma janela sem precedentes nos mecanismos do cérebro. A ressonância magnética funcional (fMRI) e a tomografia por emissão de positrões (PET) facultam insights detalhados sobre a anatomia e a função do cérebro, aperfeiçoando a precisão do diagnóstico e orientando as decisões terapêuticas.

"Os avanços nas técnicas imagiológicas, tanto moleculares como estruturais, abrem uma janela sem precedentes nos mecanismos do cérebro".

D.

A biotecnologia, ingressa neste desafio como pilar fundamental na inovação médica no que respeita à prevenção de distúrbios neurológicos. Ao desvendar os segredos do nosso código genético, a análise genómica revela predisposições a distúrbios específicos. Este conhecimento capacita os médicos a criarem estratégias de prevenção personalizadas adaptadas ao código genético único de cada indivíduo, inaugurando a era da medicina de precisão com apoio da IA. No campo da psiquiatria, a biotecnologia assume um papel de destaque possibilitando uma compreensão mais profunda da doença mental. Marcadores moleculares específicos estão a ser descobertos, dando uma nova dimensão ao diagnóstico de condições como depressão e esquizofrenia. Nesta área, a sinergia entre a IA e a biotecnologia impulsiona o desenvolvimento de marcadores moleculares específicos. Estes marcadores, nascidos da interseção da biologia molecular e da tecnologia, tornam-se ferramentas de diagnóstico essenciais, destacando as anomalias antes mesmo de se manifestarem como sintomas evidentes.

"No campo da psiquiatria, a biotecnologia assume um papel de destaque possibilitando uma compreensão mais profunda da doença mental".

Acredito que esta revolução silenciosa tem potencial para alterar o curso do diagnóstico na área da neurologia. Antevejo que o avanço destas tecnologias será futuramente determinante para lograr reduzir a severidade dos distúrbios neurológicos e prover uma franca melhoria na qualidade de vida dos pacientes. A IA emerge como uma componente primordial neste "quebra-cabeças", novos avanços computacionais urgem para que se possa tirar cada vez mais proveito desta tecnologia, nunca descurando o respeito pelos princípios éticos de acesso a dados, que são evidentemente fundamentais nesta área do diagnóstico de distúrbios neurológicos. Algoritmos de aprendizagem computacional e automática, treinados com grandes conjuntos de dados clínicos, genómicos, sinais e imagens, apoiam cada vez mais a decisão médica formulando diagnósticos diferenciais com uma rapidez e precisão sem precedentes. Convêm ressalvar que a IA amplia, ajuda, mas não substitui a perícia clínica e humana, apenas fornece evidências valiosas para orientar decisões clínicas críticas.

Pedro Miguel Rodrigues

Professor Auxiliar na Escola Superior de Biotecnologia (Católica do Porto)

D.

A EQUIPA DO DIURNA. DESEJA-LHES UMA ÓTIMA LEITURA.

Diurna.

OS TEXTOS DOS AUTORES CONVIDADOS
NÃO SÃO SUJEITOS A QUALQUER PROCESSO
DE REVISÃO, POR RESPEITO AO ESTILO
PRÓPRIO DE CADA UM.

D.

O Jornal Nacional dos Estudantes da Universidade Católica Portuguesa

Porto | Lisboa | Braga | Viseu