

# Ética, igualdade e inteligência artificial

Sofia B. Nunes<sup>1</sup>, Rui Nunes<sup>1</sup>

1. Faculdade de Medicina da Universidade do Porto, Porto, Portugal.

## Resumo

A inteligência artificial traz profundas alterações à sociedade, sejam sociais, econômicas ou culturais, com importantes repercussões em princípios éticos universalmente partilhados, como liberdade, privacidade, igualdade e não discriminação, em especial no setor da saúde. Se por um lado seu desenvolvimento pode gerar crescimento econômico, por outro pode agravar a vulnerabilidade de diversos grupos, como mulheres, idosos ou pessoas com deficiência, em decorrência de desinformação ou de vieses algorítmicos. O artigo analisa os mais relevantes instrumentos legais existentes que enquadram a inteligência artificial como ferramenta de fomento de igualdade e de inclusão geral e conclui que é preciso promover estratégias holísticas para colocar as oportunidades geradas por ela à frente dos riscos.

**Palavras-chave:** Bioética. Ética. Inteligência artificial. Direitos humanos. Políticas públicas antidiscriminatórias.

## Resumen

### Ética, igualdad e inteligencia artificial

La inteligencia artificial conlleva profundos cambios en la sociedad, ya sean sociales, económicos o culturales, con importantes repercusiones en principios éticos universalmente compartidos, como la libertad, la privacidad, la igualdad y la no discriminación, especialmente en el sector salud. Si, por un lado, su desarrollo puede generar crecimiento económico, por otro, puede agravar la vulnerabilidad de diversos grupos, como las mujeres, las personas mayores o las personas con discapacidad, a consecuencia de la desinformación o los sesgos algorítmicos. Este artículo analiza los instrumentos legales más relevantes y que enmarcan la inteligencia artificial como herramienta para promover la igualdad y la inclusión general, y concluye que es necesario promover estrategias holísticas para anteponer las oportunidades que genera a los riesgos.

**Palabras clave:** Bioética. Ética. Inteligencia artificial. Derechos humanos. Políticas públicas antidiscriminación.

## Abstract

### Ethics, equality and artificial intelligence

Artificial intelligence brings profound changes to society, whether social, economic, or cultural, with important repercussions on universally shared ethical principles such as freedom, privacy, equality, and non-discrimination, especially in the health sector. While its development can generate economic growth, it can also exacerbate the vulnerability of various groups, such as women, older adults, or people with disabilities, as a result of misinformation or algorithmic biases. This article analyzes the most relevant existing legal instruments that frame artificial intelligence as a tool for promoting equality and general inclusion and concludes that holistic strategies must be promoted to put the opportunities generated by it ahead of the risks.

**Keywords:** Bioethics. Ethics. Artificial intelligence. Human Rights. Public nondiscrimination policies.

Declararam não haver conflito de interesse.

A inteligência artificial (IA) veio revolucionar o dia a dia automatizando tarefas, aumentando a produtividade, criando novas formas de trabalho, mas também eliminando determinados empregos, com um tremendo potencial de crescimento econômico. Na medicina e restantes profissões da saúde, igualmente se antecipa essa evolução. Isso não difere do que aconteceu em revoluções tecnológicas do passado, desde a Primeira Revolução Industrial, por volta do século XVIII, com a introdução das máquinas em substituição ao trabalho manual; passando pela Segunda Revolução Industrial, no século XIX, com avanços científicos como o telégrafo e a eletrificação; a Terceira Revolução Industrial, no século XX, com a transição das indústrias para as tecnologias de informação, destacadamente os computadores e a internet; até, finalmente, a Quarta Revolução Industrial, que estamos a viver, com rápidos avanços tecnológicos. Cada uma dessas revoluções provocou profundas transformações nas sociedades e demandou que pessoas e entidades se adaptassem a uma nova realidade.

Contudo, é necessário haver medidas específicas para contrabalançar os potenciais efeitos negativos da IA à sociedade, pois, dependendo das medidas criadas, ela pode tanto criar ou intensificar desigualdades como fomentar inclusão. É essencial tornar os sistemas de IA ferramentas de inclusão, o que se faz atuando do lado criador, desde a criação da base do sistema até sua implementação, e do lado dos cidadãos, dos mais novos aos mais velhos, que precisam compreender como a IA funciona e se implementa. É, por isso, crucial apostar na educação e formação para as novas tecnologias e, assim, combater riscos para a população, como o de desinformação. Literacia em inteligência artificial, sobretudo em inteligência artificial generativa (IAGen), é essencial.

De fato, a desinformação é considerada a maior ameaça a curto prazo para qualquer sociedade e pode pôr em risco o processo democrático de um país ao gerar desconfiança e polarização e minar direitos humanos fundamentais, como o direito à saúde. Por exemplo, na União Europeia, observou-se aumento de quase 1.000% nos casos de sarampo entre 2023 e 2024 devido à desinformação acerca dos benefícios da vacinação. Na medicina, a desinformação pode ter efeito dramático ao induzir doentes a diagnósticos e tratamentos não

comprovados cientificamente. Apesar de não ser novo, é um fenômeno exacerbado pela rápida troca de informações ocasionada pelas novas tecnologias, nomeadamente as redes sociais. Se a IA pode gerar conteúdo falso sem que ninguém perceba, como se pode confiar nas informações transmitidas pelos meios de comunicação?

A legislação nacional e internacional torna-se, desse modo, fundamental para estabelecer os requisitos mínimos e as diretrizes éticas que os sistemas de IA devem seguir para que a população esteja protegida. Contudo, é necessário que as leis sejam devidamente implementadas e monitoradas, de modo a detectar mais facilmente qualquer potencial transgressão. A rápida progressão das tecnologias também exige que a legislação acompanhe seu ritmo, o que pode ser particularmente difícil; no entanto, tem de haver adaptações constantes, e a punição a cibercrimes deve ser reequacionada com vistas a salvaguardar as pessoas.

Este artigo começa por elencar os efeitos negativos do progresso da inteligência artificial; passa para as potencialidades dela, inclusive de ser motor de inclusão à luz dos princípios bioéticos universais; e termina com a legislação criada nesse âmbito tanto em nível nacional quanto internacional, com foco em Portugal e no Brasil, e nos princípios éticos específicos à IA, que devem sempre ser preponderantes.

## Inteligência artificial generativa como motor de inclusão

As revoluções tecnológicas provocam transformações nas sociedades nos mais diversos níveis, com repercussões tanto positivas como negativas, cujo peso depende das medidas implementadas. Se, por um lado, pode haver agravamento de desigualdades e exclusão social, por outro, pode-se gerar igualdade em vários eixos. O caso das novas tecnologias digitais, nomeadamente a inteligência artificial, não é diferente, mas traz novos efeitos negativos que devem ser prevenidos, os quais são visíveis nos campos social, econômico, político e cultural e afetam toda a população. Newfield<sup>1</sup> menciona seis potenciais problemas associados a sistemas de IA:

1. preconceitos sociais (sobretudo ligados a etnia e gênero);

2. opacidade (utilizadores não sabem como resultados são gerados);
3. coerção (algoritmos e dados são controlados pelos criadores e plataformas, e não pelos utilizadores);
4. violações sistêmicas de privacidade (IA depende de acesso a grandes quantidades de dados, muitas vezes de fontes diversas e que podem conter informação pessoal);
5. ausência de liberdade acadêmica no que diz respeito à investigação em empresas (objetivos comerciais podem influenciar resultados);
6. termo “inteligência” distorce o estatuto real e os efeitos desses sistemas (há potencial de utilização indevida e os próprios sistemas indicam que têm falhas).

É preciso, por isso, ter em conta as mudanças socioeconômicas que os mais diversos segmentos populacionais podem sofrer. A automação possibilitada pela IA vai ocasionar a substituição de determinadas tarefas e aumentar a produtividade, prevenindo-se crescimento econômico significativo. Reconhece-se que a IA vai tornar obsoleta uma série de empregos e revolucionar, se não todos, a maioria dos setores econômicos, como aconteceu nas revoluções tecnológicas anteriores. É esperado que, nas economias mais avançadas, haja impacto em cerca de 60% dos empregos e, nos mercados emergentes e economias menos avançadas, o impacto seja menor, de 40% e 26%, respectivamente<sup>2</sup>. Estima-se também que as mulheres vão ser mais afetadas que os homens – cerca de 21% das mulheres –, dado que ocupam empregos mais suscetíveis a tais impactos, como os dos setores da saúde e da educação, ou o trabalho administrativo<sup>3</sup>.

Não só isso, mas as capacidades cognitivas poderão ser afetadas, dado que, *embora a inteligência artificial generativa possa melhorar a eficiência dos trabalhadores, ele pode inibir o envolvimento crítico com o trabalho e potencialmente levar a uma dependência excessiva a longo prazo da ferramenta e à diminuição das competências para a resolução independente de problemas. Maior confiança na capacidade da IAGen de executar uma tarefa está relacionada a menos esforço de pensamento crítico*<sup>4</sup>. E tem de se falar também da potencial exclusão social da população mais idosa, vez que, via de regra, as gerações mais novas demonstram maior aptidão para acompanhar e adaptar-se à rápida evolução tecnológica. Seja nas interações

sociais ou no acesso a serviços, pode existir acrescida dificuldade, e até segregação, das pessoas com mais idade conforme as interações se tornam mais tecnológicas, principalmente daquelas que vivem em zonas sem internet ou com falhas nesse serviço.

A mais recente transformação digital trouxe outro grande risco para a sociedade: a rápida transmissão de informação falsa, também conhecida por desinformação e misinformação. É necessário distinguir desinformação de misinformação (traduzidas do inglês *disinformation* e *misinformation*, respectivamente). Ambas se referem a informação falsa, mas a primeira não carrega necessariamente o intuito de enganar, enquanto o segundo termo tem o objetivo de manipular. Assim, ao longo do texto, a expressão “desinformação” será usada para aludir a informação falsa, dado que não se conhecem os motivos por trás de sua transmissão.

De fato, atualmente tem havido cada vez mais desinformação, gerada com o consumo e a propagação imediata de informação nos meios de comunicação, nomeadamente as redes sociais, com potencial de abalar os sistemas democráticos. Isso não é novo, já que as campanhas de desinformação foram uma técnica bastante usada por Estados Unidos da América e União Soviética durante a Guerra Fria<sup>5</sup>. Ainda assim, a divulgação facilitada de conteúdo falso, seja intencional ou não, torna o processo de manipulação mais perigoso do que nunca. E a comunicação social tem extrema importância no sistema democrático para responsabilizar e criticar o sistema político, poder com o qual vem o dever de desmistificar potenciais informações falsas e ajudar a população a distinguir a verdade. A tecnologia não é má por si; a maneira como as pessoas a utilizam é que pode provocar danos.

O Fórum Econômico Mundial, no *Relatório de Risco Global de 2024*<sup>6</sup>, identificou a desinformação e a misinformação como o maior risco a curto prazo (dois anos) para a humanidade, salientando três aspectos:

1. Nos próximos dois anos, a misinformação e a desinformação podem perturbar radicalmente os processos eleitorais em várias economias.
2. Uma crescente desconfiança na informação, bem como nos meios de comunicação social e nos governos enquanto fontes, aprofundará a polarização nas opiniões – um ciclo vicioso que poderá desencadear agitação civil e, possivelmente, confronto.

3. Existe o risco de repressão e erosão de direitos à medida que as autoridades reprimam a proliferação de informações falsas, bem como riscos decorrentes da inação.

Não obstante, as potencialidades das novas tecnologias de informação e comunicação (TIC), impulsionadas pela IA, também são vastas e podem constituir meio de inclusão, principalmente para mulheres<sup>7-9</sup>. De fato, o acesso à educação ficou mais simples graças à disponibilidade de uma panóplia de recursos para expandir o conhecimento, desde cursos on-line a programas de ensino a distância<sup>10</sup>. Aumentou também a possibilidade de empoderamento econômico por meio da criação e do desenvolvimento de negócios em plataformas digitais e mercados on-line, que geram novas oportunidades de emprego e facilitam a conciliação entre vida profissional e pessoal, nomeadamente para mulheres, que tradicionalmente assumem maior responsabilidade nos cuidados da casa e dos filhos. A flexibilidade no trabalho – por exemplo, com o trabalho remoto – lhes proporcionou maior equilíbrio.

Ademais, as novas ferramentas digitais permitem maior envolvimento e participação política, especialmente a plataformas de redes sociais, que têm a capacidade de criar conteúdo viral, isto é, que chega a grande parte da população. No que toca à saúde e ao bem-estar, as TIC podem contribuir consideravelmente para melhorar o acesso das mulheres aos cuidados de saúde e proporcionar maior cuidado com o bem-estar.

Claro está que, para tal acontecer, é necessário mitigar os efeitos danosos, quer do lado do utilizador, quer do lado do criador da tecnologia em si. Olhando em primeiro lugar para o destinatário do conteúdo, é necessário não deixar de fora qualquer geração, desde as mais jovens até as mais velhas. Educar sobre IA, seus riscos e potencialidades é fundamental. Isso deve se dar na escola, de modo a conscientizar os jovens sobre a nova realidade, mas também em formações voltadas a pessoas mais velhas, que podem não ter tanta habilidade para lidar com as novas tecnologias. A formação ao longo da vida é indispensável para proporcionar novas competências e, assim, facilitar a adaptação à nova realidade. Existem, por exemplo, jogos on-line que estimulam de forma didática o uso de IA, como “Bad News”, que *expõe as táticas e técnicas de manipulação que são*

*usadas para enganar as pessoas e construir seguidores (...) jogar cria resistência cognitiva contra formas comuns de manipulação que o utilizador pode encontrar on-line*<sup>11</sup> e que é utilizado por alguns países na União Europeia para ensinar os mais novos a detectar notícias falsas, as ditas *fake news*.

Em Portugal, foi criado um projeto-piloto em escolas para fomentar o ensino de programação, robótica e outras temáticas relacionadas com inteligência artificial<sup>12</sup>. No Brasil, já há escolas que utilizam IA com sucesso em diversas atividades pedagógicas, mas as disparidades no acesso à tecnologia entre as escolas com mais e menos recursos (por exemplo, no que concerne a infraestruturas tecnológicas) dificultam a proliferação das potencialidades da IA entre os estudantes<sup>13</sup>.

No caso específico das mulheres, continua a existir um grande fosso no que toca à existência e participação nas áreas STEM (*science, technology, engineering and mathematics*), ou seja, nos campos mais ligados às tecnologias. Como já existem desigualdades na criação original das novas tecnologias, é esperado que depois os próprios *softwares* apresentem vieses de gênero e não sejam tão sensíveis nesse aspecto como poderiam ser se mais mulheres estivessem envolvidas em sua criação e desenvolvimento. É por isso necessário captar mais mulheres para essas áreas científicas. De fato, a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (Unesco) vem alertando para esse fato e elaborou uma série de recomendações para incorporar a perspectiva de gênero na IA, sobretudo *estabelecer uma compreensão básica dos imperativos da igualdade de gênero e de como a IA, enquanto sistema sociotécnico, pode reforçar ou desafiar a desigualdade. (...) Garantir que a igualdade de gênero seja entendida em termos de abordagem aos danos, aumentando a visibilidade das implicações de gênero da IA e incentivando as aplicações positivas da IA e os impactos benéficos no empoderamento das mulheres*<sup>14</sup>.

A IA generativa pode ser instrumento dinamizador para incluir mulheres nas áreas Stem por sua aplicação interdisciplinar, fácil uso e benefícios sociais; de acordo com as Nações Unidas, *através de políticas e iniciativas específicas destinadas a melhorar a participação e liderança das mulheres na IA Gen, tais como programas de mentoria, financiamento equitativo e defesa de reformas políticas inclusivas de gênero, podem ser feitos progressos*

significativos no sentido de alcançar os objetivos do *Objetivos de Desenvolvimento Sustentável 5 (ODS)*<sup>15</sup>, ou seja, de alcançar a igualdade de gênero e empoderar todas as mulheres e meninas<sup>16</sup>.

Apesar de haver aqui enfoque nas mulheres, tem de ser mencionado que outros grupos minoritários e especialmente vulneráveis podem sofrer desigualdades nesse âmbito. Como se demonstrou previamente, todos os cidadãos, independentemente de determinadas características, como gênero, etnia, deficiência, idade ou local de origem, podem dar válido contributo para a sociedade se devidamente integrados ao tecido social e produtivo<sup>17</sup>. Isso está em linha, aliás, com a doutrina internacional dos direitos humanos que a maioria dos países adotou<sup>18,19</sup>.

Sublinha-se a enorme interdependência entre os princípios clássicos da bioética – respeito pela autonomia, beneficência, não maleficência e justiça – e a implementação de uma inteligência artificial inclusiva e de confiança. A bioética respeita a vida e o amor pelo fenómeno vital. Se a dignidade humana é o pilar da bioética moderna, deve também se relevar a proteção de outras formas de vida, incluindo o ambiente e a biodiversidade. O próprio conceito de inclusão deve ser revisitado à luz desses princípios, porque uma sociedade só poderá ser inclusiva no futuro se for sustentável no plano econômico e ambiental.

Por outro lado, a verdadeira inclusão, ou seja, a inclusão de todos em uma sociedade moderna e avançada, pressupõe liberdade individual. Na saúde, o exercício da autonomia manifesta-se, por exemplo, no consentimento informado. A inclusão sem autonomia é desprovida de sentido, e a autonomia sem inclusão é inconsequente. Já a beneficência e a correlativa não maleficência são a base moral da ação humana. A inclusão cabal na sociedade de mulheres, idosos, pessoas com deficiência, entre outros, deve decorrer de boa intenção e de motivações humanas e solidárias, e não apenas de um contrato social. Seja este estabelecido por lei ou por convenção, é sempre de cariz mais utilitário do que benevolente.

Por seu turno, o complexo princípio bioético da justiça encontra na inclusão sua mais ampla expressão<sup>20</sup>. A ética da justiça é a ética da igualdade de oportunidades. Ou seja, igualdade em bioética é garantir a todos oportunidades sociais e econômicas efetivas para desenvolver os talentos individuais,

pelo que se deduz que só existe verdadeira igualdade de oportunidades se o acesso a saúde, educação, formação profissional ou proteção social na velhice e na doença for acompanhado, na sociedade digital contemporânea, de genuína inclusão digital. Do ponto de vista bioético, igualdade de oportunidades é instrumento para autorrealização pessoal; ou, na ótica de Amartya Sen<sup>21</sup>, para desenvolvimento do pleno potencial da pessoa humana, de suas capacidades mais avançadas, as quais só podem exprimir-se em um ambiente econômico e social promotor de inclusão.

Quando se afirma que os sistemas de IA têm de ser mais inclusivos, está-se a mencionar inclusão de todas as pessoas, independentemente de gênero, classe social, etnia, tipo de deficiência, entre outros, porque qualquer sistema de IA que não seja inclusivo pode ser bastante danoso, e os impactos mais graves incluem danos materiais aos utilizadores desses sistemas, quer se trate de más classificações de crédito injustificadas, diminuição de oportunidades de educação e emprego, diagnósticos médicos imprecisos ou detenções criminais injustificadas<sup>22</sup>.

De fato, tem-se verificado que determinados sistemas de IA, por se basearem em dados que carregam preconceitos de origem, mesmo que não intencionais, perpetuam e até intensificam as desigualdades presentes. Foi o que aconteceu, por exemplo, na empresa Amazon, que utilizou um sistema de IA no recrutamento de novos colaboradores a fim de automatizar o processo e pré-seleccioná-los e acabou por discriminar candidatas femininas, pois a ferramenta:

*foi criada em cima de padrões de currículos enviados para a empresa nos últimos dez anos. Na imensa maioria, esses currículos eram de homens, como acontece na maior parte da indústria de tecnologia, e assim considerava os candidatos homens naturalmente mais aptos para as vagas. Uma simples menção ao termo de mulheres no currículo era penalizada pela ferramenta e reduzia as chances das profissionais à vaga*<sup>23</sup>.

Ou seja, como o sistema aprendeu a fazer distinções de base sem contemplar o porquê, as discriminações sentidas no passado foram exacerbadas no presente. Por esse motivo é que o ser humano deve estar no centro dos princípios éticos de IA e integrado em todas as fases do processo

de criação e desenvolvimento desses sistemas, a fim de que haja verdadeiro respeito pela pessoa humana, independentemente de características específicas.

## Emergência de legislação relativa a inteligência artificial

Olhando agora para o lado dos criadores da tecnologia em si ou de conteúdo – por exemplo, de vídeos e imagens gerados por IA ou de *chatbots* de conversação na saúde –, é necessário que eles sigam as diretrizes estabelecidas nos regulamentos nacionais e internacionais para que haja maiores transparência e responsabilidade. Ao nível da União Europeia (UE), foi criado o Regulamento da Inteligência Artificial em 2024 – o primeiro regulamento internacional nesse âmbito –, que é baseado no risco que determinado sistema de IA gera para as pessoas e para a sociedade. Ou seja, se for contra os valores da UE, esse sistema não será permitido. A nova lei trouxe a obrigação de que as pessoas:

*que recorrem a um sistema de IA para gerar ou manipular conteúdos de imagem, áudio ou vídeo cuja semelhança considerável com pessoas, objetos, locais, entidades ou eventos reais possa levar uma pessoa a crer, erroneamente, que são autênticos ou verdadeiros (falsificações profundas/deep fakes), deverão também revelar de forma clara e perceptível que os conteúdos foram artificialmente criados ou manipulados, identificando os resultados da IA como tal e divulgando a sua origem artificial*<sup>24</sup>.

Tal regra é bastante relevante, dada a proliferação de *deep fakes*, com efeitos incrivelmente negativos na vida das pessoas por quem se fazem passar. Existem vários indivíduos famosos, como artistas e políticos, que têm sofrido esse tipo de ataque, e muitas vezes é extremamente difícil perceber se o conteúdo é real ou fictício.

Daí ser importante que as próprias plataformas tenham formas de detectar se o conteúdo foi criado artificialmente, pois, se houver intenção maliciosa por parte do criador, ele não vai respeitar quaisquer princípios. Assim, está sendo desenvolvido *software* que detecta conteúdo falso, mas será preciso

acompanhar também os avanços que acontecem na vertente dos criadores de conteúdo, ou seja, as tecnologias de ambos os lados têm de estar passo a passo. Surpreendentemente, algumas falsificações profundas/*deep fakes* ocorrem na medicina; há casos documentados de alteração deliberada de imagem de TAC ou de ressonância magnética com impacto real na saúde e na vida de pessoas.

No Brasil, foi aprovado em 2024 o Projeto de Lei 2.338, de 2023<sup>25</sup>, que regulamenta a IA e prevê igualmente uma avaliação com base no risco dos sistemas de IA. Esse documento põe:

*o ser humano como o princípio “de todas as coisas”, na centralidade das decisões (e) proíbe a avaliação de traços de personalidade e características de comportamento para prever a realização de crimes. E impede a classificação de indivíduos com base em seu comportamento social ou personalidade para determinar, de forma ilegítima e desproporcional, o acesso a bens, serviços e políticas públicas*<sup>26</sup>.

Relativamente ao conteúdo gerado de modo artificial, é bastante claro quanto à não instigação de danos e à verificação de sua autenticidade e proveniência.

As consequências de não regular sistemas de IA são imensas, e os cibercrimes que têm acontecido nos últimos tempos são prova disso, especialmente aqueles contra mulheres, com maior ocorrência de *violência sexual baseada em imagens, coloquialmente conhecida como pornografia de vingança*<sup>27</sup>. É o caso do *deep fake* de cariz sexual, o qual, dependendo do país e das relações de gênero nele existentes (por exemplo, se são mais ou menos conservadoras), pode ter repercussões desastrosas para as mulheres, como perda de emprego, renegação pela família e posteriores problemas de saúde mental.

A Diretiva da União Europeia (EU) 1.385/2024<sup>28</sup>, relativa ao combate à violência contra mulheres e à violência doméstica, vem justamente criminalizar esse tipo de crime digital. Já no Brasil, foi aprovado em 2024 um projeto de lei para punir a violência digital contra mulheres, que será inserido na Lei Maria da Penha – Lei 11.340/2006<sup>29</sup> –, a qual visa eliminar a violência doméstica e familiar contra mulheres<sup>30</sup>. Os avanços sentidos em matéria legislativa mostram a importância de acompanhar o progresso tecnológico. É, por isso, necessário que as

leis sejam devidamente implementadas e monitorizadas a fim de identificar potenciais transgressões.

Para tentar densificar os princípios éticos especificamente aplicados em uma IA de confiança, a Comissão Europeia, em 2019, criou o Grupo Independente de Peritos de Alto Nível sobre Inteligência Artificial<sup>31</sup>, que enunciou sete critérios éticos essenciais:

1. agência e supervisão humanas;
2. robustez técnica e segurança;
3. privacidade e boa governação de dados;
4. transparência;
5. diversidade, não discriminação e justiça;
6. bem-estar social e ambiental;
7. *accountability* (responsabilização).

Os critérios, de aplicação geral, pretendem garantir que a IA seja concebida e aplicada de modo a promover diversidade, não discriminação e justa distribuição dos encargos e benefícios com sua utilização. Também a Organização Mundial da Saúde (OMS) se debruçou sobre a importância de uma governança ética e inclusiva da inteligência artificial<sup>32</sup>, agora especificamente na saúde, realçando a necessidade de:

1. proteger a autonomia;
2. promover o bem-estar humano, a segurança humana e o interesse público;
3. garantir transparência, explicabilidade e inteligibilidade;
4. promover responsabilidade e prestação de contas;
5. garantir inclusão e equidade;
6. promover uma inteligência artificial responsiva e sustentável.

Foi criado o primeiro tratado internacional no nível do Conselho da Europa em 2024 – a Convenção-Quadro do Conselho da Europa sobre Inteligência Artificial e Direitos Humanos, Democracia e Estado de Direito –, que é juridicamente vinculativo e está aberto a países fora da Europa. Ele especifica os princípios éticos que todos os sistemas de IA devem seguir no respectivo ciclo de vida<sup>33</sup>:

1. dignidade humana e autonomia individual;
2. igualdade e não discriminação;
3. respeito pela privacidade e proteção de dados pessoais;
4. transparência e supervisão;
5. responsabilidade e prestação de contas;
6. confiabilidade;
7. inovação segura.

Percebe-se que existe a fundada preocupação de diferentes organismos internacionais de garantir a fiabilidade da IA e o controle e supervisão humanos, assim como assegurar a diversidade e a inclusão pela promoção de acesso e literacia digitais a grupos vulneráveis, tradicionalmente excluídos, ou pelo menos preteridos, agora com a agravante geracional. Ou seja, importa, desde já, combater o “dividendo geracional” incluindo todas as gerações e evitando interseccionalidades que agravem a exclusão.

Porém, cresce a noção de que a simples enunciação desses e outros princípios não é suficiente<sup>34</sup>. É necessário, por um lado, promover intensas campanhas de educação e formação em IA e outras competências digitais em todos os grupos etários e estratos sociais, e, por outro, implementar regulação específica, nacional e internacional, incluindo legislação que enquadre eticamente a inteligência artificial. Por exemplo, pela garantia de “igualdade epistêmica”, ou seja, de que os sistemas de IA não serão desenvolvidos de modo a privilegiar determinadas perspectivas éticas ou sistemas de conhecimento em detrimento de outros.

Tem de haver também constante evolução da legislação e das políticas públicas para que acompanhem as evoluções tecnológicas e não haja discrepância entre o progresso alcançado, com os potenciais perigos que dele advêm, e medidas defensoras do ser humano. A punição de atos discriminatórios e crimes digitais é fundamental para coibir uma atuação desumana e excessiva consequente do uso dos sistemas de IA.

Deve-se olhar para essas tecnologias também pelo aspecto positivo, isto é, por seu contributo para combater a desigualdade e a exclusão. Por exemplo, pessoas com deficiência podem e devem encontrar nesse conjunto alargado de inovações tecnológicas – que inclui a IA ou a síntese biomecânica – as ferramentas necessárias para se integrarem ao *mainstream* social<sup>35</sup>. De fato, as tecnologias foram sempre um recurso imprescindível para ultrapassar uma série de deficiências e assim permitir maior e melhor expressão da identidade pessoal e da autonomia individual. Dispositivos mecânicos ou eletrônicos permitem ultrapassar incapacidades orgânicas ou funcionais com expressão interna ou externa ao organismo humano, mas com grande valor terapêutico e enorme inclusão social<sup>36</sup>.

Em contexto médico ou de reabilitação, a IA associada a esse tipo de tecnologia poderá ser fator promotor de inclusão e igualdade<sup>37</sup>. Não se trata de atingir necessariamente um futuro pós-humano há muito desejado pela humanidade, mas de generalizar e democratizar o acesso a modernas evoluções tecnológicas tais como assistentes digitais, *software* voz-texto, avatares de língua gestual, membros protéticos avançados, exoesqueletos sofisticados, ou apoio à saúde mental<sup>38</sup>. Todas essas evoluções promovem igualdade e inclusão através de uma simbiose única e de *design* inovador e permitem nova *performance* individual pela atenuação de diferentes incapacidades.

Porém, ressalva-se que, ao falar de inclusão e igualdade, existem importantes riscos interseccionais que não apenas agravam, mas potencializam as desigualdades<sup>39</sup>, e que podem ser exacerbados pela IA se não forem detectados e corrigidos. Isto é, a combinação das já existentes desigualdades com base em gênero, etnia ou incapacidade pode intensificar vulnerabilidades previamente existentes<sup>40</sup>. Em síntese, e como refere Francis Fukuyama, *os seres humanos foram preparados pelo processo evolutivo para serem criaturas sociais que procuram naturalmente integrar-se em uma atmosfera de relacionamentos comunais*<sup>41</sup>.

A IA deve servir para aprofundar a solidariedade e a inclusão social, e não para o oposto. Na medicina e na saúde, áreas em que o uso de IA é especialmente sensível, seis princípios éticos adicionais devem ser considerados de forma sistemática para que a inteligência artificial seja considerada de confiança e, sobretudo, para que esteja sob supervisão humana e em conformidade com os objetivos do desenvolvimento sustentável: justiça, universalidade, rastreabilidade, usabilidade, robustez e explicabilidade<sup>42,43</sup>. Só assim se pode assegurar a todos o legítimo direito a um futuro aberto<sup>44</sup>.

## Considerações finais

As potencialidades da inteligência artificial são imensas, mas, para concretizá-las em benefício da humanidade e para que não se tornem riscos previsíveis, é necessário haver uma estratégia concertada para atacar várias frentes, desde o lado da concepção ao do utilizador, e, assim, fazer da IA um verdadeiro impulsionador da inclusão geral

na sociedade digital e no acesso à saúde. Contudo, o desafio de criar mecanismos específicos é complexo e exige envolvimento de diversos parceiros, sobretudo na medicina, desde entidades governamentais até empresas privadas. Programas formativos, acesso a recursos digitais para todas as pessoas, implementação e monitorização de legislação são exemplos de medidas que precisam ser implementadas para não agravar a desigualdade no novo mundo digital, incluindo no acesso à saúde.

De fato, a IA tem potencial de promover desenvolvimento econômico e bem-estar social e de se traduzir no pleno desenvolvimento humano de todos os cidadãos, em todas as sociedades. Para isso, princípios universais de bioética devem ser respeitados, sobretudo na concepção e implementação dos diversos sistemas de inteligência artificial. A presença humana (*human in the loop*) em todos os níveis de implementação da IA é a garantia do respeito à dignidade humana como pilar central da bioética contemporânea, bem como dos princípios inalienáveis de igualdade, não discriminação, responsabilidade e respeito pela privacidade, assim como de uma autonomia que contribua para a autorrealização das pessoas, e não para sua alienação. Igualdade significa que todos devem ter acesso à inclusão digital, sem discriminação por idade, gênero, renda, deficiência ou outra condição. Autonomia implica que a IA reforce a liberdade individual desde logo na saúde, permitindo escolhas mais bem informadas.

Na saúde, o exercício do direito universal de acesso a serviços de qualidade implica que as principais ameaças éticas e sociais identificadas pela generalização da IA – como preconceitos algorítmicos, exclusão digital, riscos interseccionais, desinformação e sistemáticas violações de privacidade – sejam instantaneamente detectadas, de modo que o direito à saúde seja uma realidade e que a IA contribua para o gozo efetivo dele. O recurso à inteligência artificial para otimizar a gestão hospitalar e, assim, traduzir-se em ganhos de eficiência na alocação de recursos poderá ser outra forma de melhorar o acesso à saúde e contribuir para a inclusão.

Todas essas virtualidades da IA, porém, carecem de um quadro normativo nacional e internacional para harmonizar práticas e contribuir para o bem-estar dos cidadãos. A União Europeia, por meio do Regulamento da Inteligência Artificial, aprovado

em 2024, pretende garantir que os princípios éticos e bioéticos não sejam violados. Para isso, o risco na utilização da IA é estratificado em níveis de atuação, e a privacidade individual e a proteção de dados pessoais merecem especial destaque nesse instrumento legal. O Brasil segue na mesma linha de atuação por meio do Projeto de Lei 2.338/2023, em análise no Congresso, o qual estabelece diretrizes éticas para o desenvolvimento e aplicação da IA. Esses e outros instrumentos jurídicos pretendem ressaltar não apenas os princípios bioéticos tradicionais, mas também princípios específicos de uma IA de confiança, como justiça, universalidade, rastreabilidade, usabilidade, robustez, responsabilização e a necessária explicabilidade do modo como a IA toma decisões e alcança conclusões.

Em síntese, para a IA ser considerada de confiança e um instrumento do pleno desenvolvimento humano, diversos eixos devem ser equacionados. Por exemplo, deve-se considerar que a IA tem efeitos em todos os domínios da atividade humana e influencia todos os objetivos de desenvolvimento sustentável, os quais se relacionam estreitamente com os princípios bioéticos; ou que só com a colaboração de todos os agentes envolvidos a IA pode ser verdadeiramente uma ferramenta de auxílio para os seres humanos; e, finalmente, que a igualdade como valor bioético intrínseco pode e deve ser alcançada, de sorte que cada indivíduo, recorrendo ou não à IA, desenvolva suas capacidades e alcance o pleno potencial, atingindo assim a máxima autorrealização pessoal.

## Referências

1. Newfield C. How to make “AI” intelligent; or, the question of epistemic equality. *Critical AI* [Internet]. 2023 [acesso 11 ago 2025];1(1-2). DOI: 10.1215/2834703X-10734076
2. Quase 40% dos empregos em todo o mundo serão afetados pela IA, alerta diretora do FMI. *Euronews* [Internet]. Trabalho; 15 jan 2024 [acesso 11 ago 2025]. Disponível: <https://bit.ly/42h1UIZ>
3. Will generative AI disproportionately affect the jobs of women?. *Kenan Institute of Private Enterprise* [Internet]. Kenan Insight; 18 abr 2023 [acesso 11 ago 2025]. Disponível: <https://bit.ly/4mJyW5z>
4. Lee H, Sarkar A, Tankelevitch L, Drosos I, Rintel S, Banks R, Wilson N. The impact of generative AI on critical thinking: self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers [Internet]. In: CHI '25: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems; 26 abr-1 maio; Yokohama, Japão. New York: Association for Computing Machinery; 2025 [acesso 11 ago 2025]. Tradução livre. DOI: 10.1145/3706598.3713778
5. Ward M, Pierson S, Beyer J. Formative battles: Cold War disinformation campaigns and mitigation strategies [Internet]. Washington, DC: Wilson Center; 2019 [acesso 11 ago 2025]. Disponível: <https://bit.ly/4mCi1BL>
6. World Economic Forum. *The Global Risks Report 2024* [Internet]. 19ª ed. Geneva: WEF; 2024 [acesso 11 ago 2025]. Disponível: <https://bit.ly/46uQe1w>
7. World Economic Forum. *Global Gender Gap Report 2022* [Internet]. Geneva: WEF; 2022 [acesso 11 ago 2025]. Disponível: <https://bit.ly/4mKgWYT>
8. *Estratégia de igualdade de gênero* [Internet]. Brussels: European Commission; 2020 [acesso 11 ago 2025]. Disponível: <https://bit.ly/4nVKPX0>
9. Unesco. *Unesco Priority Gender Equality Action Plan: 2014-2021* [Internet]. Paris: Unesco; 2014 [acesso 11 ago 2025]. Disponível: <https://bit.ly/3Il2RTC>
10. Nunes S. A União Europeia como guardiã dos direitos das mulheres. In: Nunes R, Rego G, Melo H, Nunes S, Duarte I, coordenadores. *Perspectives in gender equality: making education the core of public policies*. Basel: MDPI Books; 2025.
11. *Bad News* [Internet]. Groningen: Tilt; 2021 [acesso 11 ago 2025]. Tradução livre. Disponível: <https://www.getbadnews.com/en>
12. Governo vai lançar projeto-piloto para ensino de inteligência artificial e programação. *Jornal de Notícias* [Internet]. Inovação; 12 dez 2024 [acesso 11 ago 2025]. Disponível: <https://bit.ly/42iwsDX>

13. Costa AG. Escolas do Brasil já usam inteligência artificial em aulas, mas há desafios de inclusão. Terra [Internet]. Byte; 30 maio 2024 [acesso 11 ago 2025]. Disponível: <https://bit.ly/42jDkkm>
14. Unesco. Artificial intelligence and gender equality: key findings of Unesco's Global Dialogue [Internet]. Paris: Unesco; 2020 [acesso 11 ago 2025]. Tradução livre. Disponível: <https://bit.ly/4nE0f2c>
15. Nedungadi P, Ramesh M, Govindaraju V, Rao B, Berbeglia P, Raman R. Emerging leaders or persistent gaps? Generative AI research may foster women in STEM. *Int J Inf Manage* [Internet]. 2024 [acesso 11 ago 2025];77. Tradução livre. DOI: 10.1016/j.ijinfomgt.2024.102785
16. United Nations. Convention on the Elimination of All Forms of Discrimination against Women New York, 18 December 1979 [Internet]. Geneva: OHCHR; 1979 [acesso 11 ago 2025]. Disponível: <https://bit.ly/42S23MP>
17. Nunes S. Migration and economic development: the unintended consequences for technological innovation and, thus, the development of societies arising from barriers to high-skilled immigrants [dissertação]. London: University College London; 2018.
18. United Nations. United Nations Charter [Internet]. Geneva: United Nations; 1945 [acesso 11 ago 2025]. Disponível: <https://bit.ly/3Koz9xD>
19. United Nations. Universal Declaration of Human Rights [Internet]. Geneva: United Nations; 1948 [acesso 11 ago 2025]. Disponível: <https://bit.ly/46Z1QtE>
20. Nunes R. Healthcare as a universal human right: sustainability in global health. New York: Routledge; 2022.
21. Sen A. The idea of justice. Cambridge: Harvard University Press; 2009.
22. Zowghi D, Rimini F. Diversity and inclusion in artificial intelligence In: Lu Q, Zhu L, Whittle J, Xu X, editores. Responsible AI: best practices for creating trustworthy AI systems. Boston: Addison-Wesley Professional; 2023. p. 216.
23. Valeri J. Inteligência artificial utiliza base de dados que refletem preconceitos e desigualdades. *Jornal da USP* [Internet]. Atualidades; 7 jul 2023 [acesso 11 ago 2025]. Disponível: <https://bit.ly/4ml8hWC>
24. União Europeia. Regulamento 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024. Cria regras harmonizadas em matéria de inteligência artificial e que altera os Regulamentos (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e as Diretivas 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (Regulamento da Inteligência Artificial). *Jornal Oficial da União Europeia* [Internet]. Bruxelas, 12 jul 2024 [acesso 11 ago 2025]. Disponível: <https://bit.ly/4n0qT4B>
25. Projeto de Lei n° 2.338, de 2023. Dispõe sobre o uso da inteligência artificial [Internet]. Brasília: Senado Federal; 2023 [acesso 11 ago 2025]. Disponível: <https://bit.ly/474LDn3>
26. Senado aprova regulamentação da inteligência artificial; texto vai à Câmara. *Senado Notícias* [Internet]. 2024 [acesso 11 ago 2025]. Disponível: <https://bit.ly/4gKQ6ya>
27. Venema AE. Deepfakes as a security issue: why gender matters. *Women in International Security* [Internet]. 2025 [acesso 11 ago 2025]. Disponível: <https://bit.ly/4nlu73o>
28. União Europeia. Diretiva 2024/1385, de 14 de maio de 2024. Relativa ao combate à violência contra as mulheres e à violência doméstica. *Jornal Oficial da União Europeia* [Internet]. Bruxelas, 24 maio 2024 [acesso 11 ago 2025]. Disponível: <https://bit.ly/3lwYXXL>
29. Brasil. Lei n° 11.340, de 7 de agosto de 2006. Lei Maria da Penha. *Diário Oficial da União* [Internet]. Brasília, 8 ago 2006 [acesso 11 ago 2025]. Disponível: <https://bit.ly/3lMcNWg>
30. Projeto aprovado pune uso de inteligência artificial para criar material prejudicial a mulheres. *Portal da Câmara dos Deputados* [Internet]. 2024 [acesso 11 ago 2025]. Disponível: <https://bit.ly/3VDwx1c>
31. High-Level Expert Group on AI. Ethics guidelines for trustworthy AI [Internet]. Brussels: European Commission; 2019 [acesso 11 ago 2025]. Disponível: <https://bit.ly/3KLgJqM>
32. World Health Organization. Ethics & Governance of Artificial Intelligence for Health [Internet]. Geneva: WHO; 2021 [acesso 15 ago 2025]. Disponível: <https://bit.ly/3lwZ8Cp>
33. Council of Europe. Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law [Internet]. Strasbourg: Council of Europe; 2025 [acesso 11 ago 2025]. Disponível: <https://bit.ly/4nWaWx5>

34. Mittelstadt BD. Principles alone cannot guarantee ethical AI. *Nat Mach Intell* [Internet]. 2019 [acesso 11 ago 2025];1(11):501-7. Disponível: <https://bit.ly/42kUZrP>
35. Romanska M. The bionic body: disability, technology and posthumanism. *Body Space Technol* [Internet]. 2024 [acesso 11 ago 2025];23(1). DOI: 10.16995/bst.11480
36. Murray S. *Disability and the posthuman: bodies, technology, and cultural futures*. Liverpool: Liverpool University Press; 2020.
37. Shildrick M. Disability, digital technologies and the ambivalent allure of posthumanist/transhumanist futures. *Med Humanit* [Internet]. 2024 [acesso 11 ago 2025];50:685-93. DOI: 10.1136/medhum-2024-013020
38. Hayles NK. *How we became posthuman: virtual bodies in cybernetics, literature, and informatics*. Chicago: University of Chicago Press; 1999.
39. Crenshaw K. *On intersectionality: essential writings*. New York: New Press; 2019.
40. Verloo M. Multiple inequalities, intersectionality and the European Union. *Eur J Womens Stud* [Internet]. 2006 [acesso 11 ago 2025];13(3):193-210. DOI: 10.1177/1350506806065753
41. Fukuyama F. *O nosso futuro pós-humano: consequências da revolução biotecnológica*. Lisboa: Quetzal; 2002. p. 193.
42. Lekadir K, Frangi AF, Porras AR, Glocker B, Cintas C, Langlotz CP *et al*. *FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare*. *BMJ* [Internet]. 2025 [acesso 11 ago 2025];388:e081554. DOI: 10.1136/bmj-2024-081554
43. Nunes R, Nunes S. Reliable artificial intelligence: the 18th sustainable development goal. *J Ethics Leg Technol* [Internet]. 2024 [acesso 11 ago 2025];6(2):1-19. Disponível: <https://bit.ly/4nXw9qp>
44. Feinberg J. The child's right to an open future. In: Aiken W, LaFollette H, editores. *Whose child?: children's rights, parental authority and state power*. Totowa: Littlefield, Adams & Co; 1980.

Sofia B. Nunes – Doutoranda – [asnunes@med.up.pt](mailto:asnunes@med.up.pt)

 0000-0002-7182-9890

Rui Nunes – Doutor – [ruinunes@med.up.pt](mailto:ruinunes@med.up.pt)

 0000-0002-1377-9899

#### Correspondência

Sofia B. Nunes – Al. Prof. Hernâni Monteiro. CEP 4200-319. Porto, Portugal.

#### Contribuições dos autores (CRediT)

Ambos os autores contribuíram de forma igual para a concepção, elaboração, redação e revisão crítica do manuscrito, aprovaram a versão final e assumem responsabilidade pela integridade do conteúdo.

**Disponibilidade de dados:** Todos os dados utilizados ou gerados na pesquisa estão integralmente descritos e apresentados no corpo do artigo.

**Editores responsáveis:** Dilza Teresinha Ambrós Ribeiro

**Recebido:** 20.3.2025

**Revisado:** 18.8.2025

**Aprovado:** 19.8.2025