

PARECER 103_04/2024

Questões Éticas na Utilização da Inteligência Artificial no Ensino Superior

A Inteligência Artificial (IA) está a revolucionar todas as áreas da sociedade, e a educação não é exceção. Em particular, com o advento da IA generativa, estudantes e educadores passaram a ter acesso a recursos transformadores, que constituem, ao mesmo tempo, oportunidades sem precedentes e também desafios significativos.

Historicamente, a tecnologia é facilitadora de novas práticas educacionais e a Tecnologia Educacional (ou EdTech) é atualmente uma área de estudo de pleno direito. Com o surgimento dos meios audiovisuais, nos anos 60, os estudantes deixaram de estar vinculados a um local físico ou à interação direta com os educadores, podendo aceder a conteúdos educativos remotamente e/ou em formato pré-gravado¹. Mais tarde, o nascimento da indústria de semicondutores permitiu maior eficiência nas tarefas de sala de aula que, de outra forma, consumiriam muitos recursos, libertando tempo que estudantes e educadores poderiam usar para perseguir objetivos educacionais mais nucleares e complexos². Por volta de 1970, as calculadoras portáteis começaram a permitir cálculos matemáticos rápidos, enquanto a democratização da computação pessoal na década de 1980 disparou essas capacidades, com acesso a um poder de processamento cada vez maior e a ferramentas convenientes (por exemplo, processadores de texto). Mais recentemente, a Internet e tecnologias derivadas transformaram para sempre as práticas educativas; a partir da década de 1990, estudantes e educadores começaram a poder partilhar e aceder a conteúdos em uma escala sem precedentes³. Enquanto anteriormente o conhecimento fora da sala de aula estava em grande parte circunscrito a um meio físico (por exemplo, livros de biblioteca, notas de aula, ficheiros num computador local), combinada com a proliferação de dispositivos móveis (ou seja, *smartphones* e *tablets*) que se seguiram na década de 2000, a Internet permitiu que os recursos educativos se tornassem instantaneamente disponíveis para qualquer pessoa, em qualquer lugar e a partir de uma infinidade de fontes.

O panorama das ferramentas digitais a partir da década de 2020 está a contribuir ainda mais para aumentar e alargar a sala de aula, permitindo experiências mais envolventes, em linha com as últimas tendências

¹ Bos N, Groeneveld C, Bruggen J, & Brand-Gruwel S (2015). The Use of Recorded Lectures in Education and the Impact on Lecture Attendance and Exam Performance. *British Journal of Educational Technology*, 47:5, 906-917, DOI: 10.1111/bjet.12300

² Duncan DF (1983). Computers in Education are Entering the Fourth Revolution – Yet Health Education is Just Entering the Third. *Health Education*, 14:6, 72-75, DOI: 10.1080/00970050.1983.10628784

³ Eynon R (2008). The Use of the World Wide Web in Learning and Teaching in Higher Education: Reality and Rhetoric. *Innovations in Education and Teaching International*, 45:1, 15-23, DOI: 10.1080/14703290701757401

em aprendizagem baseada em projetos, sala de aula invertida e metodologias relacionadas. Estas incluem dispositivos móveis utilizados como equipamentos de laboratório/sala de aula, serviços baseados em nuvem e Internet das Coisas (IoT)⁴, mas estão-se a expandir rapidamente para abranger também tecnologias emergentes, como Realidade Virtual (VR) e Realidade Aumentada (AR).

No geral, cada avanço em EdTech trouxe vantagens em três pilares principais: (1) Acesso a recursos educacionais; (2) Eficiência do processo educativo; e (3) Personalização do percurso de aprendizagem. À luz destes pilares, e considerando como as práticas educativas se têm historicamente entrelaçado com a tecnologia, a IA é mais um trampolim no caminho evolutivo da EdTech.

Atualmente, falar de IA no Ensino Superior parece convocar de imediato o ChatGPT⁵, “ferramenta de IA generativa mais usada ou, pelo menos, a mais experimentada”⁶, ainda que estejamos perante uma paisagem bastante mais ampla e que ainda não vislumbramos nem na plenitude nem na totalidade.

Naturalmente, o assunto está na ordem do dia⁷ e várias organizações internacionais já se pronunciaram sobre questões éticas, nomeadamente a UNESCO com “Recommendation on the Ethics of Artificial Intelligence”⁸ em 2023, ou a Organização Mundial da Saúde com “Ethics and governance of artificial intelligence for health: guidance on large multi-modal models”⁹ em 2024. Anteriormente, também o Parlamento Europeu se pronunciou, com o estudo “The ethics of artificial intelligence: Issues and initiatives”¹⁰ em 2020.

Evidentemente, a comunidade académica nacional tem estado a estudar o assunto e, desde 2023, têm sido realizadas formações e conferências sobre as temáticas da “Inteligência Artificial no Ensino Superior”¹¹, “Inteligência Artificial na Educação - Ameaças e oportunidades para o processo ensino-aprendizagem”¹², “Inteligência Artificial generativa na educação”¹³, “A Inteligência Artificial e o seu impacto no ensino-aprendizagem”¹⁴ ou “Ferramentas de IA no ensino, aprendizagem e avaliação”¹⁵, entre outros; no Congresso

⁴ Pinto J, Silva HP, Melo F, & Fred A (2020). ScientIST: Biomedical Engineering Experiments Supported by Mobile Devices, Cloud and IoT. *Signals*, 1:2, 110-120, DOI: 10.3390/signals1020006

⁵ “O Chat GPT é um modelo de linguagem que permite que as pessoas interajam com um computador de forma mais natural e conversacional. GPT significa “Transformador pré-treinado generativo” e é o nome dado a uma família de modelos de linguagem natural desenvolvidos por uma Inteligência Artificial. Também é conhecido como uma forma de IA generativa por sua capacidade de produzir resultados originais.” UNESCO (2023). ChatGPT e inteligência artificial na educação superior: guia de início rápido. Em https://unesdoc.unesco.org/ark:/48223/pf0000385146_por

⁶ “O ChatGPT é a ferramenta de IA generativa mais usada ou, pelo menos, a mais experimentada. Sabemos da importância das perguntas na qualidade das respostas obtida. O resultado de um prompt ao ChatGPT é tanto mais qualificado quanto mais culto for o utilizador” Oliveira, L. & Pinto, M. (2023). A Inteligência Artificial na Educação - Ameaças e oportunidades para o processo ensino-aprendizagem. P.6. Em <http://hdl.handle.net/10400.22/22779>

⁷ Seminário Instituto Politécnico de Viseu, abril 2024,

⁸ UNESCO (2023). Recommendation on the Ethics of Artificial Intelligence. Disponível em <https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence>

⁹ WHO (2024). Ethics and governance of artificial intelligence for health: guidance on large multi-modal models. Em <https://www.who.int/publications/i/item/9789240084759>

¹⁰ European Parliament. (2020). The ethics of artificial intelligence: Issues and initiatives. Panel for the Future of Science and Technology. Em [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU\(2020\)634452_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf)

¹¹ Alusão a um ciclo de conferências na Universidade de Aveiro, em março de 2023. Em <https://www.ua.pt/pt/noticias/4/80173>

¹² Evento realizado em abril de 2023 na Escola Superior de Media Artes e Design (ESMAD) do Politécnico do Porto (P.PORTO). Livro organizado por Lino Oliveira e Mário Pinto em <http://hdl.handle.net/10400.22/22779>.

¹³ Referência a formação realizada na UAlg em novembro de 2023, <https://esec.uaig.pt/ualg-organiza-formacao-em-inteligencia-artificial-generativa-na-educacao>

¹⁴ Universidade Aberta, ciclo de conferências, Institutional Response to Generative AI in Learning, Teaching and Assessing. 7 dezembro de 2023. <https://portal.uab.pt/a-inteligencia-artificial-e-o-seu-impacto-no-ensino-aprendizagem/>

¹⁵ Universidade de Lisboa. Instituto Superior Técnico. 9 de janeiro de 2024. Foi apresentado o relatório do trabalho da Comissão REFlet, bem como o Documento da Deliberação do CP sobre utilização de IA nos métodos de ensino e de avaliação de conhecimentos. Em <https://www.ulisboa.pt/evento/programa-contigo-6a-edicao-ferramentas-de-ia-no-ensino-aprendizagem-e-avaliacao>

Nacional de Práticas Pedagógicas de 2023, o assunto foi matéria de conferência e de diversas comunicações livres.

Alguns conselhos nacionais de ética e/ou bioética¹⁶ abordaram as preocupações éticas da IA, sendo obviamente mais salientes as referências na área das ciências da vida e da saúde. Em maio de 2024, o Conselho Nacional de Ética para as Ciências da Vida apresentou o Livro Branco "Inteligência Artificial (IA): inquietações sociais, propostas éticas e orientações políticas"¹⁷.

Com este parecer, pretende a CE-IPS partilhar alguns princípios e recomendações, que possam servir de apoio nesta fase do desenvolvimento do conhecimento. Procedemos a um breve enquadramento em ordem a suportar as recomendações.

Enquadramento

O que se entende por IA “O conceito de **Inteligência Artificial** (IA) aplica-se a sistemas que apresentam um comportamento inteligente, analisando o seu ambiente e tomando medidas— com um determinado nível de autonomia — para atingir objetivos específicos. Os sistemas baseados em inteligência artificial podem ser puramente confinados ao software, atuando no mundo virtual (por exemplo, assistentes de voz, programas de análise de imagens, motores de busca, sistemas de reconhecimento facial e de discurso), ou podem ser integrados em dispositivos físicos (por exemplo, robôs avançados, automóveis autónomos, veículos aéreos não tripulados ou aplicações da Internet das coisas)”¹⁸.

Ou “**Inteligência Artificial** (IA) refere-se à capacidade de algoritmos integrados em sistemas e ferramentas de aprender com os dados para que possam executar tarefas automatizadas sem programação explícita de cada passo por um ser humano. IA generativa é uma categoria de técnicas de IA em que algoritmos são treinados em conjuntos de dados que podem ser usados para gerar novos conteúdos, como texto, imagens ou vídeo. Estas orientações abordam um tipo de IA generativa, grandes modelos multimodais (LMMs), que podem aceitar um ou mais tipos de entrada de dados e gerar diversos resultados que não estão limitados ao tipo de dados inseridos no algoritmo.”¹⁹.

Coloquialmente, o termo **IA** é aplicado a um sistema que imita funções cognitivas como aprendizagem, raciocínio, e resolução de problemas. A **IA generativa** é uma técnica que usa algoritmos de aprendizagem de máquina para criar novos dados a partir de um conjunto existente. Estas técnicas podem ter várias aplicações, incluindo geração de dados, análise de dados, ou criação de simulações realistas de sistemas complexos.

Utilizações da IA “Tal como aconteceu no passado com o motor a vapor ou a eletricidade, a IA está a transformar o nosso mundo, a nossa sociedade e a nossa indústria. O aumento da

¹⁶ Por exemplo, o Conselho de Ética alemão – Deutscher Ethikrat (2023). Humans and Machines – Challenges of Artificial Intelligence. Em <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf>

¹⁷ CNECV (2024). Livro Branco. Inteligência Artificial (IA): inquietações sociais, propostas éticas e orientações políticas. <https://www.cneqv.pt/pt/publicacoes/monografias/livro-branco-inteligencia-artificial-ai-inquietacoes-sociais-pro>

¹⁸ Comunicação da Comissão Europeia. [Inteligência artificial para a Europa](#). Abril 2018. p. 1.

¹⁹ Ethics and governance of artificial intelligence for health: guidance on large multi-modal models. Em <https://www.who.int/publications/i/item/9789240084759>

capacidade computacional, a disponibilidade de dados e a evolução dos algoritmos fizeram da IA uma das tecnologias mais estratégicas do século XXI²⁰.

São muitos os exemplos do “que a IA é capaz de fazer em todos os setores”²¹; foi previsto amplo uso e aplicação nos cuidados de saúde, na investigação científica, na saúde pública e no desenvolvimento de medicamentos. E também na educação e no ensino superior.

Como se desenvolve IA

“A aprendizagem profunda foi um ponto de viragem para a IA, com um enorme avanço na execução de tarefas específicas, como o reconhecimento de voz e de imagem, ou a tradução automática. O treino de um algoritmo de aprendizagem profunda para fins de classificação de objetos consiste na sua exposição a um grande volume de exemplos etiquetados (por exemplo, imagens), que estão categorizados de forma correta (por exemplo, imagens de aviões). Uma vez treinados, os algoritmos são capazes de classificar corretamente objetos que nunca viram, em alguns casos, com uma exatidão superior à dos seres humanos. Foram alcançados progressos significativos no âmbito destas tecnologias mediante a utilização de grandes conjuntos de dados e de uma capacidade computacional sem precedentes.”²²

“O desenvolvimento da IA requer um grande volume de dados. A aprendizagem automática, um tipo de IA, opera mediante a identificação de padrões nos dados disponíveis, aplicando em seguida o conhecimento adquirido a novos dados. Quanto maior for a amostra, mais fácil será descobrir relações entre os dados, mesmo as mais subtis. No que respeita à utilização da IA, os ambientes mais ricos em dados também oferecem mais oportunidades, visto que os algoritmos aprendem e interagem com o seu ambiente por intermédio dos dados.”²³

Existe virtualmente uma infinidade de recursos disponíveis para aprofundar este tema, dos quais se destacam o livro “Deep Learning”²⁴ ou os currículos “Artificial Intelligence for Beginners”²⁵ e “Generative AI for Beginners”²⁶.

Desafios referidos a nível europeu quanto à IA

Considerando o documento **Inteligência Artificial para a Europa**, a União Europeia pugna por uma abordagem sustentável às tecnologias, a mudança com base nos valores da União, “com a garantia simultânea de um elevado nível de proteção dos dados, direitos digitais e normas éticas”²⁷. “Em termos globais, existem três **grandes desafios para a UE**, que salientam o papel fundamental da educação e da formação, incluindo dos próprios professores e formadores, cuja responsabilidade cabe aos Estados-Membros.

- (1) O primeiro desafio consiste em preparar a sociedade como um todo. Tal significa ajudar todos os cidadãos a desenvolverem competências digitais básicas, assim como competências complementares que não podem ser substituídas por nenhuma máquina, como o pensamento crítico, a criatividade ou as capacidades de gestão. Articula-se com “não deixar ninguém para trás”.
- (2) Em segundo lugar, a UE precisa de concentrar os seus esforços para ajudar aqueles que ocupam postos de trabalho mais suscetíveis de serem transformados ou de desaparecer devido à automatização, à robótica e à inteligência digital.

²⁰ Comunicação da Comissão Europeia. [Inteligência artificial para a Europa](#). Abril 2018.

²¹ [Inteligência artificial para a Europa](#) “Na Dinamarca, a IA está a ajudar a salvar vidas, permitindo aos serviços de emergência diagnosticar paragens cardíacas ou outros problemas de saúde com base na voz do autor de uma chamada. Na Áustria, está a ajudar os radiologistas a detetar tumores com maior exatidão, comparando instantaneamente os resultados dos exames de raios X com um elevado volume de outros dados clínicos. Muitas explorações agrícolas de toda a Europa já estão a utilizar a IA para controlar a circulação, a temperatura e o consumo de alimentos dos animais. O sistema de IA pode depois adaptar o aquecimento e os equipamentos para alimentação automaticamente, para ajudar os agricultores a monitorizar o bem-estar dos animais e libertando-os para outras tarefas. A IA também está a ajudar os fabricantes europeus a tornarem-se mais eficientes, o que contribui para que as fábricas regressem à Europa.”

²² [Inteligência artificial para a Europa](#), p. 11.

²³ Idem, p. 11.

²⁴ Goodfellow I, Bengio Y, and Courville A (2016). “Deep Learning”. MIT Press. Em <https://www.deeplearningbook.org/>

²⁵ Artificial Intelligence for Beginners- A Curriculum. Em <https://microsoft.github.io/ai-for-beginners/>

²⁶ Generative AI for Beginners (Version 2)- A Course. Em <https://microsoft.github.io/generative-ai-for-beginners/#/>

²⁷ Conselho Europeu de outubro de 2017. <http://data.consilium.europa.eu/doc/document/ST-14-2017-INIT/pt/pdf>

Trata-se também de garantir o acesso de todos os cidadãos à proteção social, em conformidade com o Pilar Europeu dos Direitos Sociais.

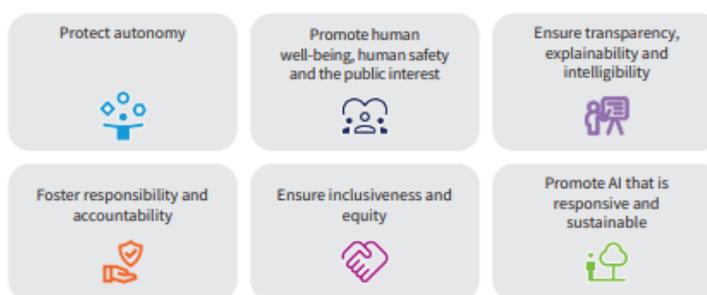
- (3) Por fim, a UE necessita de formar mais especialistas em IA, com base na sua longa tradição de excelência académica, de criar o ambiente adequado para que estes possam trabalhar na UE e de atrair mais talentos vindos do estrangeiro. Portanto, estimular o talento, a diversidade e a interdisciplinaridade

Previsto um projeto de orientações para as questões de ética da IA, tendo em devida consideração a Carta dos Direitos Fundamentais da União Europeia.

Em 2021, a OMS publicou orientações abrangentes sobre a **ética e a governação da IA para a saúde**. Os princípios devem orientar o desenvolvimento e a implantação da IA nos cuidados de saúde por uma vasta gama de partes interessadas, incluindo governos, agências do setor público, investigadores, empresas e implementadores.

Os princípios são: (1) proteger a autonomia; (2) promover o bem-estar humano, a segurança humana e o interesse público; (3) garantir transparência, “explicabilidade” e inteligibilidade; (4) fomentar a responsabilidade e a prestação de contas; (5) garantir a inclusão e a equidade; e (6) promover uma IA que seja responsiva e sustentável.

Figure 1: WHO consensus ethical principles for use of AI for health



Diretrizes para Literacia em IA de acordo com o EU ACT

De acordo com a proposta de regulamento do Artigo 4.º-B do EU ACT1 acerca da **Literacia em IA**:

“1. Quando o presente regulamento for aplicado, a União e os Estados-Membros devem promover medidas para o desenvolvimento de um nível suficiente de literacia em IA, em todos os setores e tendo em conta as diferentes necessidades dos grupos de fornecedores, responsáveis pela implantação e pessoas afetadas, nomeadamente através da educação e da formação, de programas de qualificação e de requalificação e garantindo, ao mesmo tempo, um equilíbrio adequado de género e de idade, com vista a permitir um controlo democrático dos sistemas de IA.”

“2. Os fornecedores e os responsáveis pela implantação de sistemas de IA devem adotar medidas para garantir um nível suficiente de literacia em IA do seu pessoal e de outras pessoas envolvidas na operação e utilização de sistemas de IA em seu nome, tendo em conta os seus conhecimentos técnicos, experiência, qualificações académicas e formação e o contexto em que os sistemas de IA vão ser utilizados, bem como as pessoas ou grupos de pessoas relativamente às quais os sistemas de IA vão ser utilizados.”

“3. Essas medidas de literacia devem consistir, em especial, no ensino de noções e competências básicas sobre os sistemas de IA e o seu funcionamento, incluindo os diferentes tipos de produtos e utilizações e os respetivos riscos e benefícios.”²⁸

²⁸ Alterações aprovadas pelo Parlamento Europeu, em 14 de junho de 2023, sobre a proposta de regulamento do Parlamento Europeu e do Conselho que estabelece regras harmonizadas em matéria de inteligência artificial (Regulamento Inteligência Artificial) e altera determinados atos legislativos da união (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))

Assim, verificamos que a proposta demonstra e instrui para uma clara inserção da literacia em IA nos programas educativos e outras iniciativas (como ações de formação, cursos breves, palestras). As IES deverão desenvolver programas específicos de IA para os seus estudantes, tal como garantir que todos obtêm o mesmo nível de formação em IA. Uma forma muito clara de atingir esta necessidade será introduzir nas suas ofertas formativas (independente da área científica de cada curso) uma unidade curricular acerca da IA. Para além da oferta da formação as IES deverão também garantir que os estudantes compreendam os benefícios e os aspetos negativos dos sistemas que utilizam IA.

IA e Ensino Superior (ES)

No “Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators”²⁹ (CE, 2022) distinguem entre ferramentas de IA focadas em ensinar estudantes e ferramentas de IA focadas em apoiar os estudantes. Em termos de apoio aos estudantes, uma das vantagens mais significativas da IA é o potencial que oferece para abordagens de aprendizagem personalizadas. A aprendizagem personalizada refere-se à capacidade de adaptar aspetos da aprendizagem processo, tal conteúdo ou ritmo de acordo com habilidades, interesses e preferências de aprendizagem individuais.

No relatório “**Teaching with AI – Assessment, Feedback and Personalisation**” refere-se que uma questão fundamental no campo da IA na educação diz respeito a justiça, responsabilidade, transparência e ética (fairness, accountability, transparency, and ethics - FATE) de ferramentas EdTech baseadas em algoritmos, como os sistemas tutores inteligentes ou sistemas de avaliação automática.

Utilizações Emergentes da IA no Ensino Superior

Os domínios de aplicação mais salientes de EdTech baseada em IA podem ser resumidos como: (A) Linguagem; (B) Audiovisual; (C) Desenvolvimento; e (D) Aprendizagem.

No que diz respeito à Linguagem (A), as ferramentas existentes são utilizadas principalmente para redação, revisão e geração de texto. Exemplos incluem o Grammarly³⁰, que tem uma integração perfeita com software desktop, dispositivos móveis e web, permitindo verificação ortográfica, edição e até mesmo geração de conteúdo em tempo real. O ChatGPT³¹ conquistou o mundo³², ao trazer Large Language Models (LLMs) para as massas; tornou-se uma ferramenta indispensável para redação (ou seja, melhorar textos com base em rascunhos preliminares apresentado pelo utilizador), além das capacidades avançadas de geração de conteúdo com base em *prompts* (isto é “perguntas” inseridas pelos utilizadores). No que diz respeito ao Audiovisual (B), a IA está a ser cada vez mais utilizada para superar as limitações de direitos de autor e para agilizar a criação de conteúdos. Estudantes e educadores muitas vezes enfrentam problemas com direitos de autor ao produzir conteúdo, especialmente imagens; ferramentas como Midjourney³³ ou Dall-E³⁴ foram pioneiras na capacidade de criar imagens originais a partir de *prompts* baseados em texto. Na verdade, a criação de conteúdo baseada em IA já se estende também ao domínio de áudio e vídeo; ferramentas como o Google Vertex AI Studio³⁵ podem ser usadas para criar fala a partir de texto, enquanto o Sora³⁶ pode ser usado para criar vídeos envolventes a partir de *prompts*

²⁹ European Commission, Directorate-General for Education, Youth, Sport and Culture, Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, Publications Office of the European Union, 2022, <https://data.europa.eu/doi/10.2766/153756>

³⁰ <https://www.grammarly.com/>

³¹ <https://chat.openai.com/>

³² Editorial (2023). ChatGPT is a Black Box: How AI Research can Break it Open. Nature, 619:7971, 671-672, DOI: 10.1038/d41586-023-02366-2

³³ <https://www.midjourney.com/home>

³⁴ <https://openai.com/dall-e-3>

³⁵ <https://console.cloud.google.com/vertex-ai/generative>

³⁶ <https://openai.com/sora>

definidos pelo utilizador. A tecnologia neste domínio chega ao ponto de permitir filmagens narrativas realistas (por exemplo, Synthesia³⁷).

O Desenvolvimento (C), por outro lado, é bastante abrangente, variando desde a produção de materiais educacionais assistida por IA até atividades ainda mais especializadas (por exemplo, programação). O Microsoft Copilot é um assistente abrangente para vários aplicativos que pode auxiliar com bastante competência na cocriação; um exemplo poderoso é a integração com o PowerPoint³⁸, que pode ser usado para redigir apresentações em aspectos tão simples como o índice ou tão sofisticados como toda a apresentação de slides. GitHub Copilot³⁹ é um exemplo de assistente que pode ser usado para programação na maioria das linguagens convencionais.

Por último, a IA é uma ferramenta inestimável para apoiar o processo de Aprendizagem (D), nomeadamente no que diz respeito ao estudo, avaliação e aconselhamento. Ferramentas como o Jasper Text Summarizer⁴⁰ são cada vez mais utilizadas para resumir artigos científicos, auxiliando caminhos de aprendizagem personalizados de forma mais eficiente. A avaliação e a supervisão, tanto autónomas como num ambiente educacional mais formal, também estão em ascensão; embora ainda faltem ferramentas suficientemente consolidadas, o Gradescope⁴¹ é um exemplo representativo de uma ferramenta de classificação assistida por IA. Indiscutivelmente, um dos usos mais difundidos da IA na educação é no aconselhamento (ou como “companheiro de estudo”)⁴²; LLMs como o ChatGPT podem efetivamente constituir um tutor sempre disponível, treinado virtualmente com muito mais conhecimento (tanto em termos de amplitude quanto de profundidade) do que os seus equivalentes humanos.

Riscos da IA no Ensino Superior

“A proteção dos direitos humanos e da dignidade é a pedra angular da Recomendação, baseada no avanço de princípios fundamentais como a transparência e a equidade, lembrando sempre a importância da supervisão humana dos sistemas de IA”⁴³.

As rápidas mudanças também suscitam profundas preocupações éticas. Estas surgem do potencial que os sistemas de IA têm para discriminação, incorporar preconceitos, contribuir para a degradação climática, para ameaçar os direitos humanos e muito mais. Esses riscos associados à IA já começaram a agravar-se, somando-se às desigualdades existentes, resultando em danos adicionais para grupos já marginalizados.

Outro aspeto crítico está relacionado com a avaliação em sala de aula; em particular, os relatórios escritos têm sido historicamente um componente de avaliação essencial em várias unidades curriculares, e ainda mais em projectos finais de curso, teses e dissertações. Devido à crescente sofisticação dos LLMs modernos, converge-se para um cenário em que a maior parte do conteúdo escrito será em grande parte assistido por IA ou totalmente produzido por IA⁴⁴. Isto é ainda mais preocupante à luz da falta de ferramentas eficazes para detectar conteúdos fabricados.

³⁷ <https://www.synthesia.io/>

³⁸ <https://support.microsoft.com/en-us/copilot-powerpoint>

³⁹ <https://visualstudio.microsoft.com/github-copilot/>

⁴⁰ <https://www.jasper.ai/tools/text-summarizer>

⁴¹ <https://www.gradescope.com/>

⁴² Favero TG (2024). Using Artificial Intelligence Platforms to Support Student Learning in Physiology. *Advances in Physiology Education*, 48:2, 193-199, DOI: 10.1152/advan.00213.2023

⁴³ UNESCO, 2023. Ethics of Artificial Intelligence. “The protection of human rights and dignity is the cornerstone of the Recommendation, based on the advancement of fundamental principles such as transparency and fairness, always remembering the importance of human oversight of AI systems. However, what makes the Recommendation exceptionally applicable are its extensive Policy Action Areas, which allow policymakers to translate the core values and principles into action with respect to data governance, environment and ecosystems, gender, education and research, and health and social wellbeing, among many other spheres”

⁴⁴ Kolade O, Owoseni A, & Egbetokun A (2024). Is AI Changing Learning and Assessment as We Know It? Evidence from a ChatGPT Experiment and Conceptual Framework. *Heliyon*, 10:4, e25953, DOI: 10.1016/j.heliyon.2024.e25953

É muito evidente que a IA irá gradualmente superar os humanos na maioria das tarefas sistemáticas e analíticas; a educação não é exceção. Contudo, os sistemas de IA não são perfeitos, e é conhecida a tendência dos sistemas para na ausência de melhores respostas gerarem conteúdo fictício (vulgarmente designada por “alucinações”).

Carta Portuguesa de Direitos Humanos na Era Digital, 2021

Consideremos a “Carta Portuguesa de Direitos Humanos na Era Digital”⁴⁵, que “as normas que na ordem jurídica portuguesa consagram e tutelam direitos, liberdades e garantias são plenamente aplicáveis no ciberespaço” (artigo 2º). Visa “assegurar um ambiente digital que fomente e defenda os direitos humanos”.

No artigo 9º, relativo a inteligência artificial e robôs, determina que “A utilização da inteligência artificial deve ser orientada pelo respeito dos direitos fundamentais, garantindo um justo equilíbrio entre os princípios da explicabilidade, da segurança, da transparência e da responsabilidade, que atenda às circunstâncias de cada caso concreto e estabeleça processos destinados a evitar quaisquer preconceitos e formas de discriminação”. Mais especifica que “As decisões com impacto significativo na esfera dos destinatários que sejam tomadas mediante o uso de algoritmos devem ser comunicadas aos interessados, sendo suscetíveis de recurso e auditáveis, nos termos previstos na lei.”

Também são aplicáveis “à criação e ao uso de robôs os princípios da beneficência, da não-maleficência, do respeito pela autonomia humana e pela justiça, bem como os princípios e valores consagrados no artigo 2.º do Tratado da União Europeia, designadamente a não discriminação e a tolerância.”

IES em Portugal

Algumas Instituições de ES, em Portugal, já se pronunciaram do ponto de vista pedagógico e científico. Foi o caso do Instituto Superior Técnico - defendeu a que a IA deve ser utilizada por professores e estudantes, assim como as ferramentas de IA generativa, como o ChatGPT, considerando que lhes cabe um papel pioneiro na “exploração da aplicação das ferramentas de IA, de que os modelos generativos de linguagem são exemplos, nas mais variadas áreas, nomeadamente na formação dos nossos estudantes para metodologias de trabalho emergentes em engenharia, ciência e tecnologia”⁴⁶.

Daqui decorreu que não deverá ser adotada nenhuma medida geral de proibição no uso de ferramentas que usem IA nos processos de ensino ou em métodos de avaliação. Sendo que os elementos de avaliação, com acesso a ferramentas de IA, mais personalizados e adaptativos, devem ser combinados para robustez e fiabilidade com elementos de avaliação tradicionais, como exames/testes. Preferencialmente, devem juntar-se a componentes em que haja interação direta e efetiva professor-estudantes, como orais, de forma a aferir a efetiva aquisição de conhecimentos e de competências.

Síntese das preocupações éticas

As tecnologias de IA podem ser de grande utilidade mas também suscitam preocupações éticas fundamentais,

- i. por exemplo em relação os preconceitos que podem incorporar e exacerbar, potencialmente resultando em discriminação, desigualdade, exclusão e ameaça à diversidade cultural, social e biológica e divisões sociais ou económicas;
- ii. a necessidade de transparência e compreensibilidade do funcionamento dos algoritmos e dos dados com que foram treinados;
- iii. o impacto potencial sobre a dignidade humana, direitos humanos e liberdades fundamentais, bem como sobre processos sociais, económicos, políticos e culturais.

⁴⁵ Lei n.º 27/2021 Diário da República n.º 95/2021, Série I de 2021-05-17.

⁴⁶ O documento que a comissão REFlet elaborou está disponível em <https://conselhopedagogico.tecnico.ulisboa.pt/informacoes>.

Suscitam-se questões sobre quando e como a Inteligência Artificial poderia, e deveria, ser usada para a tomada de decisões autônomas, sendo que um sistema de IA não pode substituir a responsabilidade e a prestação de contas humanas.

Os aspetos éticos têm sido categorizados em três grupos: prestação de contas (não discriminatória e justa); responsabilidade (agência humana e supervisão); e transparência (robustez técnica e governança de dados).

Numa *scoping review*, identificou-se que no âmbito dos estudos de investigação abrangido, o ímpeto contemporâneo para envolver o público em considerações éticas relacionadas com a IA gira predominantemente em torno de duas motivações principais: inovação e legitimação. “Abordar as opiniões do público sobre os desafios éticos da IA também pode ser usado como uma forma de prevenção de riscos para reduzir conflitos e encerrar debates vitais em áreas de contenção”⁴⁷. A agência humana e a supervisão emergiram como a dimensão ética predominante, ou seja, o foco da centralidade humana, que é progressivamente integrada nos discursos públicos sobre IA. A seguir, a questão ética mais frequentemente discutida foi a privacidade e a governação de dados. O foco em formas seguras de armazenamento e proteção de dados (desde a entrada até os dados de saída), o teste e a certificação de produtos e serviços habilitados para IA, os riscos de violações de dados exigem que se encontre um equilíbrio entre o alargamento o acesso aos dados e a garantia da confidencialidade e do respeito à privacidade. Diversidade, não discriminação e justiça, bem-estar social e ambiental, robustez técnica e segurança, transparência e responsabilização foram as questões éticas menos frequentemente discutidas nos estudos incluídos nesta revisão – também parecer ser aquelas para as quais soluções técnicas podem ser ou já foram desenvolvidas. A transparência ganha alta relevância nas diretrizes éticas porque este princípio se tornou uma condição pró-ética, para explicabilidade e divulgação.

Recomendações

Considerando que

- a. existe um profundo e dinâmico impacto da inteligência artificial nas sociedades, meio ambiente, ecossistemas e vidas humanas, reconhecendo-se aceleração no uso da informação e comunicação digital, com recurso a tecnologias de IA,
- b. a IA se difundiu muito rapidamente, se desenvolve velozmente, com diversas técnicas entre as quais a IA generativa,
- c. as tecnologias de IA podem ser de grande utilidade mas também suscitam preocupações éticas fundamentais,
- d. alguns dos principais desafios para a responsabilização estão relacionados com discriminação e preconceitos, pois os algoritmos de IA são treinados em grandes fontes de dados e podem ter vieses inerentes ou introduzidos nas regras de treino,

⁴⁷ Machado, H. ; Silva, S.& Neiva, L. (2023). Publics' views on ethical challenges of artificial intelligence: a scoping review. *AI and Ethics*. <https://doi.org/10.1007/s43681-023-00387-1>

- e. os níveis de credibilidade e confiabilidade aparentes, actualmente transmitidos pelas ferramentas de IA, podem condicionar os processos de aprendizagem, nomeadamente no que diz respeito à relação entre os diferentes agentes do processo de ensino,
- f. importa ponderar as vantagens e/ou benefícios de atender à relação com o mundo do trabalho, pois quem usa melhor e de forma mais eficiente as ferramentas de IA, poderá vir a ter vantagem competitiva na sua integração em tarefas qualificadas nas diversas áreas
- g. se acredita que a IA tem o potencial de transformar as práticas administrativas, pela automação⁴⁸, e as práticas pedagógicas, tornando-as mais personalizadas, com ambientes de aprendizagem adaptativa, mais eficazes e eficientes⁴⁹,

emitimos as seguintes recomendações:

1. **No IPS, deve prevalecer a orientação da conduta pelos princípios e valores éticos.** O respeito pela dignidade humana, a proteção e promoção dos direitos humanos e das liberdades fundamentais são essenciais. Importa reiterar os valores fundamentais e princípios éticos, expressos no [Código de Ética e Conduta](#) e que o Instituto Politécnico de Setúbal adota e promove.
2. **O desenvolvimento e aplicação das tecnologias tem de incluir uma abordagem centrada nas pessoas,** consciente do contexto ético e cultural dos utilizadores, e respeitando os seus direitos, incluindo os direitos em plataformas digitais⁵⁰, cibersegurança, direito à liberdade de criação e à proteção dos conteúdos⁵¹ bem como direito à proteção contra a geolocalização abusiva⁵²
3. **Quem utiliza as tecnologias de IA, seja no ensino ou na investigação, tem de se assegurar da devida proteção dos direitos humanos e que o uso é proporcional ao objetivo e apropriado ao**

⁴⁸ Referência a automação das tarefas administrativas- a IA pode ajudar a reduzir o tempo gasto em tarefas administrativas, permitindo que os professores se concentrem mais no ensino e na interação com os estudantes. A IA pode automatizar tarefas como a correção de testes, atribuição de notas e criação de relatórios, libertando tempo para os professores se concentrarem no desenvolvimento de estratégias de ensino eficazes. A IA pode ajudar a analisar grandes quantidades de dados educacionais, permitindo que os educadores identifiquem padrões e tendências que possam ajudá-los a melhorar suas práticas pedagógicas. Cf. Oliveira, L. & Pinto, M. (2023). A Inteligência Artificial na Educação- Ameaças e oportunidades para o processo ensino-aprendizagem. P.6. Em <http://hdl.handle.net/10400.22/22779>

⁴⁹ Incluindo aqui que a IA generativa pode ser usada para traduzir automaticamente o conteúdo das aulas em diferentes idiomas e/ou criar material multimídia

⁵⁰ Carta Portuguesa de Direitos Humanos na Era Digital, 2021. Artigo 14.º “1. Na utilização de plataformas digitais, todos têm o direito de: a) Receber informação clara e simples sobre as condições de prestação de serviços quando utilizem plataformas que viabilizam fluxos de informação e comunicação; b) Exercer nessas plataformas os direitos garantidos pela presente Carta e na demais legislação aplicável; c) Ver garantida a proteção do seu perfil, incluindo a sua recuperação se necessário, bem como de obter cópia dos dados pessoais que lhes digam respeito nos termos previstos na lei; d) Apresentar reclamações e recorrer a meios alternativos de resolução de conflitos nos termos previstos na lei. 2- O Estado promove a utilização pelas plataformas digitais de sinaléticas gráficas que transmitam de forma clara e simples a política de privacidade que asseguram aos seus utilizadores.”

⁵¹ Idem, artigo 16º. “1- Todos têm direito à livre criação intelectual, artística, científica e técnica, bem como a beneficiarem, no ambiente digital, da proteção legalmente conferida às obras, prestações, produções e outros conteúdos protegidos por direitos de propriedade intelectual. 2- As medidas proporcionais, adequadas e eficazes com vista a impedir o acesso ou a remover conteúdos disponibilizados em manifesta violação do direito de autor e direitos conexos são objeto de lei especial.”

⁵² Idem, artigo 17º. “Todos têm direito à proteção contra a recolha e o tratamento ilegais de informação sobre a sua localização quando efetuem uma chamada obtida a partir de qualquer equipamento. 2- A utilização dos dados da posição geográfica do equipamento de um utilizador só pode ser feita com o seu consentimento ou autorização legal.”

contexto. Por isso, a escolha de usar sistemas de IA e qual método de IA usar deve ser fundamentada - verificando que o método escolhido é apropriado e proporcional para atingir um determinado objetivo, que a utilização da IA não infringe os direitos humanos e é baseada em rigorosos fundamentos científicos. Os intervenientes na utilização da IA devem promover a justiça social e a equidade, numa abordagem inclusiva de forma que os benefícios do uso sejam acessíveis a todos os envolvidos, tendo em conta as necessidades específicas das pessoas, em particular se em situação de vulnerabilidade.

4. **Não seja adotada nenhuma medida geral de proibição no uso de ferramentas que usem IA nos processos de ensino ou em métodos de avaliação.** Mais ainda, estudantes e educadores deverão ser incentivados a utilizar ativamente ferramentas de IA, uma vez que as mesmas representam grandes vantagens competitivas para todos os agentes educativos.
5. **Haja valorização e investimento na inserção da literacia em IA nos planos de estudos,** eventualmente com unidades curriculares transversais ou com os formatos que forem mais adequados aos cursos e contextos.
6. **De forma prudente, sejam estabelecidas regras para a utilização das ferramentas IA,** num equilíbrio que permita testar e desenvolver os talentos dos estudantes. Por isso, também se releva explicitar os recursos autorizados, ou seja, em qualquer prova de avaliação, deverão ser especificados os recursos (incluindo ferramentas de IA) que os estudantes podem recorrer legitimamente, e isto em cada unidade curricular. E ao nível do curso o que é aceitável como limite de utilização.
7. **Sempre que forem utilizadas ferramentas de IA para a edição ou produção de conteúdos, deverão ser devidamente referenciadas** quais as ferramentas utilizadas e qual o papel que desempenharam - se os estudantes utilizarem para gerar imagens, proof reading, sínteses, etc... podem colocar a informação em nota de rodapé ou nos acknowledgements identificando os recursos utilizados para escrever a secção X ou Y.
8. **Os métodos de avaliação devem ser reapreciados** pois importa que, sempre que possível, exista interação direta e efetiva entre o docente e o estudante, de forma a aferir a efetiva aquisição de conhecimentos e de competências. Também se recomenda que os métodos de avaliação fortemente dependentes de componentes escritas sejam adaptados e/ou complementados de modo a promover uma escrita mais concisa e/ou outros meios escritos mais enraizados no pensamento criativo que permitam aos estudantes maior eficácia na persecução dos objetivos educativos.
9. **Para promover o envolvimento contínuo dos estudantes, importa reforçar a aprendizagem didática com pensamento crítico e aprendizagem experiencial.** O mundo real é inerentemente interativo e, em grande medida, prático e humano. Isto aumenta a importância das

abordagens de “aprender fazendo”, especialmente quando se tratam tópicos que requerem uma interface com o mundo físico, que pode ser assistido pela IA, mas atualmente não é facilmente mediado apenas pela IA.

10. **Deve proceder-se a implementação de testes e procedimentos de validação das informações que estão na base de processos de tomada de decisão**, promovendo um uso apropriado e eficaz das tecnologias, uma vez que algumas ferramentas de IA têm um grande potencial para a desinformação. Será necessária uma reflexão cuidadosa para uma utilização responsável e para evitar resultados problemáticos e danos. É necessário reconhecer com seriedade o risco de utilização das ferramentas da IA como substitutas do juízo individual, pelo que importa promover o raciocínio e juízo crítico.
11. **Os utilizadores da IA envidem todos os esforços para minimizar e evitar a perpetuação de práticas discriminatórias ou aplicações e resultados tendenciosos**. Reforça-se a necessidade de utilizar fontes de informações pluralistas e confiáveis, considerando os esforços para mitigar riscos tais como disseminação de conteúdos inventados, *fake news*, discursos de ódio e outras formas danosas de uso indevido de dados.
12. **Os dados dos sistemas de IA são recolhidos, utilizados, partilhados, arquivados e eliminados de forma consistente com o respeito pela privacidade e proteção dos dados**. Também se tem discutido a necessidade de considerar a cedência dos direitos de autoria para o utilizador, responsabilizando-o pela validação e robustez da informação e resultados. Notamos que a IA ainda é uma ferramenta, sendo a última responsabilidade e prestação de contas das pessoas que a utilizam.

Referências bibliográficas

- Bos N, Groeneveld C, Bruggen J, & Brand-Gruwel S (2015). The Use of Recorded Lectures in Education and the Impact on Lecture Attendance and Exam Performance. *British Journal of Educational Technology*, 47:5, 906-917, DOI: 10.1111/bjet.12300
- Comissão Europeia (2018). [Inteligência artificial para a Europa](#). Comunicação da Comissão ao Parlamento Europeu, ao Conselho Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões.
- Conselho Europeu (2017). Reunião do Conselho Europeu (19 de outubro de 2017) – Conclusões. <http://data.consilium.europa.eu/doc/document/ST-14-2017-INIT/pt/pdf>
- Deutscher Ethikrat (2023). Humans and Machines – Challenges of Artificial Intelligence. Em <https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf>
- Duncan DF (1983). Computers in Education are Entering the Fourth Revolution – Yet Health Education is Just Entering the Third. *Health Education*, 14:6, 72-75, DOI: 10.1080/00970050.1983.10628784
- Editorial (2023). ChatGPT is a Black Box: How AI Research can Break it Open. *Nature*, 619:7971, 671-672, DOI: 10.1038/d41586-023-02366-2

- European Commission (2022). Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, Publications Office of the European Union. <https://data.europa.eu/doi/10.2766/153756>
- European Parliament (2020). The ethics of artificial intelligence: Issues and initiatives. Panel for the Future of Science and Technology. Em [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU\(2020\)634452_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf)
- Eynon R (2008). The Use of the World Wide Web in Learning and Teaching in Higher Education: Reality and Rhetoric. *Innovations in Education and Teaching International*, 45:1, 15-23, DOI: 10.1080/14703290701757401
- Favero TG (2024). Using Artificial Intelligence Platforms to Support Student Learning in Physiology. *Advances in Physiology Education*, 48:2, 193-199, DOI: 10.1152/advan.00213.2023
- Goodfellow I, Bengio Y, and Courville A (2016). "Deep Learning". MIT Press. Em <https://www.deeplearningbook.org/>
- Instituto Superior Técnico (2024). Documento da Deliberação do CP sobre utilização de IA nos métodos de ensino e de avaliação de conhecimentos. Em <https://www.ulisboa.pt/evento/programa-contigo-6a-edicao-ferramentas-de-ia-no-ensino-aprendizagem-e-avaliacao>
- Kolade O, Owoseni A, & Egbetokun A (2024). Is AI Changing Learning and Assessment as We Know It? Evidence from a ChatGPT Experiment and Conceptual Framework. *Heliyon*, 10:4, e25953, DOI: 10.1016/j.heliyon.2024.e25953
- Lei n.º 27/2021 Diário da República n.º 95, Série I, 17 de maio. Carta Portuguesa de Direitos Humanos na Era Digital
- Machado, H. ; Silva, S.& Neiva, L. (2023). Publics' views on ethical challenges of artificial intelligence: a scoping review. *AI and Ethics*. <https://doi.org/10.1007/s43681-023-00387-1>
- Oliveira, L. & Pinto, M. (2023). A Inteligência Artificial na Educação- Ameaças e oportunidades para o processo ensino-aprendizagem. Em <http://hdl.handle.net/10400.22/22779>
- Pinto J, Silva HP, Melo F, & Fred A (2020). ScientIST: Biomedical Engineering Experiments Supported by Mobile Devices, Cloud and IoT. *Signals*, 1:2, 110-120, DOI: 10.3390/signals1020006
- UNESCO (2023). ChatGPT e inteligência artificial na educação superior: guia de início rápido. Em https://unesdoc.unesco.org/ark:/48223/pf0000385146_por
- UNESCO (2023). Recommendation on the Ethics of Artificial Intelligence. Disponível em <https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence>
- WHO (2024). Ethics and governance of artificial intelligence for health: guidance on large multi-modal models. Em <https://www.who.int/publications/i/item/9789240084759>

Relatores: Hugo Plácido da Silva, Luís Leitão, Lucília Nunes

Apresentação e discussão na CE na reunião plenária de 22 de abril de 2024.

Apresentação no Seminário de Práticas Pedagógicas do IPS, a 23 de maio de 2024.

Versão presente aprovada por unanimidade na reunião plenária de 17 de junho de 2024.

Presidente da CE-IPS