

Guidelines para uma Utilização Responsável da Inteligência Artificial Generativa no Ensino Superior e na Publicação Científica¹

Comissão de Ética FMUP/RISE-Health

Rui Nunes (Presidente), Guilhermina Rego (Vice-Presidente), Cristina Prudêncio, Helena Santos Leitão, Inês Falcão Pires, Ivone Duarte, Jorge Luiz dos Santos, Mónica Correia, Natália Oliva Teles, Patrícia Ribeiro, Rita Negrão, Francisca Rego (Coordenadora Sec. Executivo)

26 de junho de 2025

I- Introdução

A utilização da inteligência artificial generativa (IAGenerativa) no ensino, na produção de conhecimento e na publicação de resultados é seguramente uma das áreas que mais necessita de literacia aprofundada, de modo a garantir elevados níveis de integridade e de compromisso académico (Stahl et al. 2024). Regras estritas devem existir para que estudantes, docentes e cientistas saibam como utilizar com ética e integridade a inteligência artificial generativa em todo o seu percurso académico. A IAGenerativa é já hoje de utilização recorrente no meio académico e tem o potencial de melhorar significativamente os processos de aprendizagem e a aquisição de competências nos mais diferentes domínios científicos. Existindo mesmo um dever genérico da comunidade académica de promover a capacitação de estudantes e docentes na aquisição de competências estruturais no domínio da inteligência artificial, definindo claramente quais devem ser essas competências e o modo como devem ser adquiridas (OCDE 2025).

1

Ante o exposto, a utilização responsável da Inteligência artificial deve estar sustentada em três eixos essenciais: a) no respeito por princípios éticos, valores morais e códigos de conduta; b) na existência de um propósito educativo e/ou científico claro; e c) na complementaridade em relação aos processos de ensino-aprendizagem.

Algumas das formas positivas e aceitáveis de usar a IAGenerativa incluem:

1. *Brainstorming* de ideias por meio de *prompts*.
2. Ajuda no raciocínio lógico e obtenção de explicações sobre perguntas e conceitos difíceis, incluindo nos *prompts* as principais referências da temática em questão para obter respostas mais fidedignas.

¹ Proposta elaborada no Centro de Bioética da Faculdade de Medicina da Universidade do Porto e adotada pela Comissão de Ética da FMUP/RISE-Health. Relatores: Rui Nunes e Sofia B. Nunes.

3. Auxílio a superar o bloqueio criativo por meio do diálogo interativo com a IAGenerativa.
4. Auto-tutoria por meio de conversação com a IAGenerativa.
5. Aulas colaborativas entre docentes e estudantes favorecendo uma aprendizagem ativa com a partilha de *prompts*.
6. Criação de questões práticas e autotestes.
7. Organização e resumo de anotações pessoais.
8. Planejamento e auxílio na estrutura de um trabalho.
9. Ajuda na gramática, ortografia, escrita e tradução de textos.
10. Resumo de um texto, artigo ou livro (tendo em atenção os direitos de autor).

Aquilo que se exige é uma utilização com ética e integridade da inteligência artificial generativa não sendo penalizável a sua utilização no estrito respeito por linhas diretrizes consensuais na matéria. Daí a importância da discussão e consensualização na comunidade académica de guidelines essenciais para uma utilização ética e responsável da IAGenerativa no ensino superior. Estas *Guidelines para uma Utilização Responsável da Inteligência Artificial Generativa no Ensino Superior e na Publicação Científica* foram apresentadas e discutidas em uma conferência de consenso realizada na Faculdade de Medicina da Universidade do Porto no dia 30 de abril de 2025, sendo de realçar a especial relevância de esta temática no ensino médico (Ahuja 2019, Goh et al. 2025).

No entanto, o uso de IAGenerativa acarreta uma série de riscos e limitações que podem impactar a qualidade do trabalho e respetiva avaliação. Embora seja legítimo utilizar a IAGenerativa para auxiliar na aprendizagem, pode ser necessário, por exemplo, em uma prova ou outra avaliação, demonstrar as competências e habilidades sem o auxílio de essas ferramentas.

2

Embora a IAGenerativa possa ser útil para algumas tarefas, é essencial que se conheçam as suas limitações, incluindo as seguintes:

1. As ferramentas de IAGenerativa são sistemas de linguagem, e não bancos de dados de conhecimento. Funcionam prevendo a próxima palavra, imagem ou trecho plausível, através de códigos de programação e recorrendo a padrões que foram “aprendidos” em grandes conjuntos de dados (dados de treino). (Gottlieb et al. 2023). É importante conhecer em pormenor os processos iterativos da IA, isto é o refinamento dos sistemas de IA através de múltiplos ciclos de desenho, testagem e melhoramento.
2. A IAGenerativa não tem compreensão do que gera, ainda que comece a denotar alguma agência, mesmo intencionalidade, na resposta aos *prompts*. Os resultados totalmente gerados por IA podem frequentemente carecer de raciocínio crítico e profundidade de expressão e compreensão (Gottlieb et al. 2023b).

3. A IAGenerativa é genericamente considerada “amoral”, não sabendo que é errado gerar conteúdo ofensivo, impreciso ou enganoso, podendo mesmo contribuir para a desinformação. Porém, versões recentes da IAGenerativa denotam já a influência acrescida de um conjunto de valores éticos nucleares (Nuñez 2025). A geração de conhecimento através da IAGenerativa pode disseminar informação falsa contribuindo para a desinformação da sociedade a propósito do conhecimento científico e dos resultados da pesquisa (Baker 2016). Por exemplo, no caso das vacinas ou hábitos alimentares. Garantir a verificabilidade, veracidade e reprodutibilidade da informação é fundamental.
4. Os conjuntos de dados com os quais a IAGenerativa aprende têm falhas e contêm imprecisões, vieses e limitações. A resposta ao *prompt* pode incluir vieses, distorções (alucinações), deturpações e plágio, dependentes não apenas do tipo de algoritmo (que tende a ser melhorado com as novas versões), mas sobretudo dos dados de treino. Um ser humano com conhecimento apropriado deve supervisionar o *output* identificando alucinações e imprecisões.
5. A IAGenerativa gera textos que nem sempre são factualmente corretos podendo produzir citações e referências falsas. Pode criar *software*/código com falhas de segurança e bugs. Frequentemente, o código ou cálculo produzido pela IA parecerá plausível, mas contém erros. A IAGenerativa não consegue ainda uniformizar adequadamente padrões de imagens. Os dados de treino da IAGenerativa não estão necessariamente atualizados, podendo ter significativas limitações.
6. É possível potenciar-se a fraude académica e científica, nomeadamente de plágio, pela reprodução de matéria criada por outros autores. Como é também provável a retenção da informação pelo *chatbot* (*prompts* e conteúdos) podendo ser reutilizada em resposta a outros autores. Autores, editores, e revisores devem estar cientes da possibilidade de aumento do risco de plágio com a sua utilização (Appel et al. 2023). Os direitos autorais e a prevenção de plágio são alguns dos principais dilemas éticos da IA. De facto, a IAGenerativa pode incluir “plágio oculto” – o que significa utilizar palavras e ideias de autores humanos sem os referenciar. A IAGenerativa pode usar bibliotecas ilegais, e o material gerado pela IA pode infringir direitos de autor ou de propriedade intelectual.
7. O trabalho escrito é uma forma fundamental de demonstrar pensamento crítico. Confiar apenas em resultados gerados por IA impedirá que se desenvolva as habilidades adquiridas ao fazê-lo sozinho através da escrita manual. Ou seja, é essencial aprender a pensar e a construir um argumento por meio da escrita manual. A IAGenerativa não substitui isso.

8. É importante ressaltar que a dependência excessiva de ferramentas de IA simplesmente para gerar conteúdo escrito, imagem, código de software ou análise reduz a oportunidade de estudantes e docentes desenvolverem e praticarem habilidades essenciais (por exemplo, escrita, pensamento crítico, avaliação, análise, codificação e raciocínio). Todos esses são aspetos importantes na universidade e na vida profissional (Yu et al 2021).
9. A quebra de confidencialidade e a falta de proteção da privacidade de trabalho original não publicado, bem como a partilha de dados pessoais dos investigadores pelas empresas de IA Generativa a terceiras partes, mesmo com fins comerciais é uma possibilidade real (Thorbecke 2023). O que suscita importantes questões éticas relacionadas com a segurança e proteção dos dados e com o inevitável recurso dos cidadãos ao seu legítimo “direito ao esquecimento”, ou seja, o direito a solicitar o apagamento dos seus dados pessoais dos sistemas digitais (Correia et al 2021). Daí que se recorra progressivamente a dados sintéticos, dado que não são provenientes de nenhuma pessoa (Kokosi 2022).
10. A rápida evolução para sistemas totalmente autónomos com falta de compreensão do seu método de atuação (*black box*) pela falta de explicabilidade da IA deve ser acompanhada (Ioannidis J. 2018). Podendo originar falta de efetiva supervisão humana da sua atuação e da necessária replicabilidade e reprodutibilidade da pesquisa científica e, assim, da confirmação da sua validade (National Academies of Sciences 2019) A rápida evolução dos robôs humanoides totalmente autónomos será um dos grandes desafios da humanidade pela enorme capacidade intelectual que irão deter e pelo nível de agência e de intencionalidade potenciais.
11. Não é de excluir a possibilidade de atribuição de personalidade jurídica à IA, e mesmo de personalidade moral (Jones 2025), originando uma diluição de responsabilidade pela utilização das modernas tecnologias digitais. Sendo fundamental começar a desenhar um novo quadro jurídico de atribuição de responsabilidade pelas consequências da utilização da inteligência artificial generativa, de um modo proporcional ao seu nível de autonomia.
12. É imperativa a necessidade de regulação legal internacional, por exemplo através da generalização do *AI Act*, o Regulamento da União Europeia sobre Inteligência Artificial (European Commission 2024), incluindo o impacto da IA Generativa na sustentabilidade ambiental e na Uma Só Saúde/*One Health* (Akbarialiabad et al. 2025).

13. A possibilidade de agravar as desigualdades sociais – a nível de uma comunidade ou mesmo no plano global contribuindo para a exclusão social – inclui novas formas de desigualdade de cariz intergeracional dado que as gerações mais velhas não têm o mesmo tipo de acesso à literacia digital em IA (EU Health Policy Platform 2023). Daí a importância do controlo e supervisão do fluxo de informação (Lee et al. 2023).
14. O estímulo para o isolamento social dos utilizadores com desenvolvimento de empatia e afetividade pelos sistemas de IA com sérias repercussões no plano da interação social e da solidariedade entre as gerações. Sobretudo a “geração beta” (2025-2039), com elevados níveis de imersão nas novas tecnologias digitais e uma elevada dependência das tecnologias digitais, será especialmente afetada pela IA sendo previsível forte impacto na evolução da civilização humana. (Nunes et al. 2024).

Para além de estas limitações da inteligência artificial generativa como ferramenta em si, deve prevenir-se a má conduta académica, seja intencional ou por mero desconhecimento das normas de conduta, através de ações de formação abrangentes e regulares. Alguns exemplos de má conduta académica que devem ser do conhecimento geral incluem o plágio, o autoplágio (ou seja, submeter o mesmo trabalho para avaliação duas vezes na mesma instituição ou em instituições diferentes, ou o mesmo artigo para publicação em duas revistas sem mencionar expressamente a publicação prévia), a falsificação, fraude contratual (em que um estudante ou docente paga para que um trabalho seja escrito ou editado por outra pessoa), ou engano/personificação (ou seja, personificar outro estudante ou permitir que outra pessoa se faça passar por um estudante em uma avaliação) (Baker 2016).

5

Neste contexto sugere-se um conjunto de Guidelines para uma Utilização Responsável da Inteligência Artificial Generativa no Ensino Superior e, seguidamente, na Publicação Científica de acordo com as normas internacionalmente estabelecidas (World Association of Medical Editors. 2025).

II- Guidelines para uma Utilização Responsável da Inteligência Artificial Generativa no Ensino Superior

Ao usar qualquer ferramenta IAGenerativa no Ensino Superior, algumas guidelines devem ser de aplicação generalizada, nomeadamente:

1. É legítimo recorrer à IAGenerativa para auxiliar a aprendizagem quando necessário, bem como para melhorar ou corrigir um texto escrito, mas nunca se deve copiar e colar nenhum resultado da IAGenerativa no próprio trabalho avaliado. Fazer isso constitui má conduta académica. Todos os trabalhos submetidos para avaliação devem ser originais. Deve existir uma declaração específica para o efeito.

2. Em caso de dúvidas deve consultar-se sempre o diretor de curso/responsável pela unidade curricular sobre o uso de IAGenerativa no trabalho avaliado. Alguns trabalhos avaliados podem incentivar o uso de IAGenerativa, enquanto outros podem impor restrições. A IAGenerativa pode ser um valioso recurso de aprendizagem, mas nunca deve ser usada como substituto do trabalho avaliado. Deve evitar-se o risco de dependência excessiva da IAGenerativa. Os docentes devem dinamizar aulas que promovam um equilíbrio entre a realização de trabalhos com e sem recurso à IA. A referência apresentada nos trabalhos escritos deve ser verificada e valorizada.
3. Deve dar-se o devido crédito ao uso das ferramentas de IA. Antes de entregar o trabalho para avaliação, deve reconhecer-se o uso da IAGenerativa, quando utilizada. Trata-se de um sinal de transparência no que respeita à sua utilização e ao conteúdo gerado. Não é aceitável apresentar conteúdo gerado por IA como se fosse trabalho individual. Este reconhecimento deve incluir:
 - a) Nome, versão, editora e URL do sistema de IAGenerativa utilizado. Por exemplo ChatGPT, DeepSeek, Claude, Gemini, ou DALL-E.
 - b) Breve descrição (frase única) do contexto em que a ferramenta foi utilizada.
 - c) Partilha dos *prompts* usados para gerar material a partir de uma ferramenta de IAGenerativa.
 - d) O resultado obtido ou um extrato do resultado obtido.
 - e) Como o resultado foi usado, editado ou incorporado a um trabalho.
4. Deve proteger-se os dados pessoais evitando enviar dados pessoais para um sistema de IAGenerativa.
5. Deve respeitar-se os direitos de autor não se podendo enviar materiais protegidos por direitos autorais para uma plataforma IAGenerativa sem a autorização do detentor de esses direitos (valorizando-se, assim, a originalidade a paternidade intelectual).
6. Deve verificar-se sempre a precisão das informações produzidas pela IAGenerativa, incluindo referências e citações.
7. Devem diversificar-se as fontes não confiando exclusivamente na IAGenerativa devendo esta complementar, mas não substituir, as fontes tradicionais.

III- Guidelines para uma Utilização Responsável da Inteligência Artificial Generativa na Publicação Científica

1. A IAGenerativa não pode ser um autor devendo os autores ser transparentes quando recorrem à IAGenerativa fornecendo informações sobre como foi utilizada.
2. Os autores que enviarem um artigo no qual a IAGenerativa foi usada para redigir um novo texto devem anotar tal uso nos agradecimentos (acknowledgments). Todos os prompts usados para gerar novo texto ou para converter texto ou prompts de texto em tabelas ou ilustrações devem ser especificados.
3. Quando a IAGenerativa é usada para realizar ou gerar um trabalho analítico, ajudar a relatar resultados (por exemplo gerando tabelas ou figuras), ou escrever códigos de computador, ou efetuar imagens, isso deve ser declarado no corpo do artigo, tanto no resumo como na secção dos métodos e desenho do estudo.
4. Para permitir o escrutínio científico, incluindo a replicação e a identificação de falsificações, devem ser fornecidas indicações completas sobre os prompts utilizados para gerar os resultados da investigação, bem como a ferramenta de IA usada e a sua versão.
5. Os autores são responsáveis pelo material fornecido por IAGenerativa para o seu artigo (incluindo a exatidão do que é apresentado e a ausência de plágio) e pelo reconhecimento adequado de todas as fontes (incluindo as fontes originais do material gerado por IAGenerativa).
6. Os editores e revisores não devem recorrer à IAGenerativa por razões de confidencialidade e de proteção de dados da informação obtida para avaliação.
7. Os editores devem dispor de ferramentas apropriadas para ajudá-los a detetar conteúdos gerados ou alterados por IA. Essas ferramentas devem ser disponibilizadas aos editores para benefício da ciência e do público em geral, e para ajudar a garantir a integridade científica e académica.

Referências:

- Ahuja AS. 2019. The impact of artificial intelligence in medicine on the future role of the physician. PeerJ, Life & Environment, 10 <https://doi.org/10.7717/peerj.7702>
- Akbarialiabad H, et al. 2025. The utility of generative AI in advancing global health. NEJM AI February 10. DOI: 10.1056/AIp2400875

- Appel G, et al. 2023. Generative AI has an intellectual property problem. April 07, *Harvard Business Review*. <https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem>
- Baker M. 2016. Is there a reproducibility crisis? *Nature* 533 (7604): 452-454.
- Correia M, Rego G, Nunes R. 2021. Gender transition: Is there a right to be forgotten. *Health Care Analysis* 29: 283-300.
- EU Health Policy Platform. 2023. Navigating Health Inequalities in the EU through Artificial Intelligence. EU Health Policy Platform Thematic Network 2022-23: Joint Statement Presentation 19th April 2023, London.
- European Commission. 2024. AI Act. Brussels. <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
- Goh E. et al. 2025. GPT-4 assistance for improvement of physician performance on patient care tasks: A randomized controlled trial. *Nature Medicine*. <https://doi.org/10.1038/s41591-024-03456-y>
- Gottlieb M, et al. 2023. ChatGPT and conversational artificial intelligence: friend, foe, or future of research? *American Journal of Emergency Medicine* 70: 81-83. <https://doi.org/10.1016/j.ajem.2023.05.018>.
- Gottlieb M, et al. 2023b. ChatGPT and conversational artificial intelligence: Ethics in the eye of the beholder. *American Journal of Emergency Medicine* 70 (2023) 191
- Ioannidis J. 2018. Why replication has more scientific value than original discovery. *Behavioral and Brain Sciences* 41 e137 <http://www.nature.com/articles/s41562-016-0021>
- International Committee of Medical Journals Editors, 2025. Who is an author? Defining the role of authors and contributors. ICMJE. <https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html#two>
- Jones N. 2025. How should we test AI for human level intelligence? OpenAI's o3 electrifies quest. *Nature News* 14 January 2025.
- Kokosi T, Harron K. 2022. Synthetic data in medical research. *BMJ Medicine* 1: e000167. DOI:10.1136/bmjmed-2022-000167
- Lee P, Bubeck S, Petro J. 2023. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *New England Journal of Medicine* 388 (13): 1233-1239.
- National Academies of Sciences. 2019. Reproducibility and replicability in science. National Academies Press (US); Washington (DC): May 7. ISBN-13: 978-0-309-48616-3
- Nunes R, Nunes, SB. 2024. Reliable Artificial Intelligence: The 18th Sustainable Development Goal. *Journal of Ethics and Legal Technologies*, 6(2), December 2024.

- Nuñez M. 2025. Anthropic just analyzed 700.000 Claude conversations and found its AI has a moral code of its own. VentureBeat. <https://venturebeat.com/ai/anthropic-just-analyzed-700000-claude-conversations-and-found-its-ai-has-a-moral-code-of-its-own/>
- OCDE: Empowering Learners for the Age of AI. An AI Literacy Framework for Primary and Secondary Education. European Commission and Organization for Economic Cooperation and Development (OECD), Brussels, May 2025.
- Stahl B, Eke D. 2024. The ethics of ChatGPT – Exploring the ethical issues of an emerging technology. International Journal of Information Management 74 (2024) 102700
- Thorbecke C. 2023. Don't tell anything to a chatbot you want to keep private. CNN Business. April 6. <https://www.cnn.com/2023/04/06/tech/chatgpt-ai-privacy-concerns/index.html>
- Yu H, Zhou Z. 2021. Optimization of IoT-Based artificial intelligence assisted telemedicine health analysis system. IEEE Access VOLUME 9, 85035-85048
- World Association of Medical Editors. 2025. Chatbots, ChatGPT, and scholarly manuscripts: WAME recommendations on ChatGPT and chatbots in relation to scholarly publications. Available at: <https://wame.org/page3.php?id=106>.