

Linhas de orientação para a Utilização Responsável de Inteligência Artificial Generativa na Investigação

Porto, abril 2025



INESCTEC

**CREATING A FULFILLING
AND SUSTAINABLE FUTURE
THROUGH IMPACTFUL
SCIENCE, TECHNOLOGY
AND INNOVATION.**

Introdução

A Inteligência Artificial Generativa (IAGen) traduz-se em diversas oportunidades, abrangendo diferentes setores. A investigação é uma das áreas que poderá ser significativamente influenciada pela IAGen - nomeadamente, através da agilização das descobertas científicas e da melhoria dos processos de investigação e de validação, em termos de eficácia e de celeridade. Assim, os investigadores devem tirar partido das potencialidades da IAGen, mas também manter uma abordagem crítica no que concerne a utilização dos resultados produzidos, cientes da finalidade, capacidades e limitações das ferramentas.

O objetivo geral da IAGen para produção do texto é gerar a sequência de palavras mais provável com base no treino que recebeu - por norma, com uma componente de aleatoriedade que acaba por dificultar a replicabilidade. Determinadas ferramentas, comumente denominadas RAG (*Retrieval-Augmented Generative AI*), acrescentam ao mecanismo de treino base etapas complementares de recuperação de informação, raciocínio e adequação cultural. No entanto, existem vários casos de resultados inadequados após a utilização de ferramentas IAGen em atividades científicas, como viés, alucinações e imprecisões.

O impacto da IAGen na investigação e nos processos científicos exige, portanto, reflexão e ponderação; por exemplo, ao resumir artigos, gerar hipóteses, debater ou explorar ideias, analisar dados, redigir ou traduzir. Em muitos aspetos e contextos, o uso destas ferramentas pode prejudicar a integridade da investigação e levantar questões sobre possíveis práticas científicas incorretas e a disseminação de desinformação. Alguns riscos estão associados às limitações técnicas das próprias ferramentas, sendo que outros decorrem do uso (intencional ou não) da ferramenta de formas que prejudicam as práticas de investigação. Como tal, é importante que os investigadores estejam cientes dos potenciais benefícios e barreiras da utilização de IAGen. Estas orientações visam prevenir o uso indevido e garantir que este tipo de soluções desempenha um papel positivo na melhoria das práticas de investigação.

Para tal, e tendo como base as orientações desenvolvidas pelo *European Research Area Forum*¹ - complementando-as com uma visão geral das orientações existentes em instituições que atuam nas mesmas áreas -, o INESC TEC propõe um conjunto de orientações para a utilização da IAGen na investigação, de acordo com princípios éticos fundamentais, para benefício da sua comunidade. Tendo em conta o ritmo acelerado dos desenvolvimentos tecnológicos neste domínio, estas orientações devem ser vistas como um conjunto de linhas orientadoras gerais e dinâmicas, ajudando os investigadores a avaliar e a refletir sobre possíveis implicações éticas relacionadas com a privacidade ou outros aspetos, decorrentes da utilização da IAGen nos seus trabalhos de investigação. Estas orientações devem ser acompanhadas pela promoção da participação individual e coletiva de todos os agentes envolvidos em atividades de investigação, a fim de desenvolver uma cultura adequada e eficaz de utilização responsável destas ferramentas. As iniciativas de literacia e de formação em IA irão desempenhar, certamente, um papel decisivo no reforço das competências neste domínio.

¹ Estas orientações são uma adaptação das *Living guidelines on the responsible use of generative AI in research* - Comissão Europeia, Direção-Geral da Investigação e Inovação, © União Europeia, 2024, licenciadas sob a licença *Creative Commons Attribution 4.0 International* (CC BY 4.0). Outras fontes e referências obrigatórias incluem o Código Europeu de Conduta para a Integridade na Investigação, o trabalho e as orientações sobre IA confiável do *High-Level Expert Group on AI*, bem como o Código de Ética do INESC TEC.

Linhas de orientação para a Utilização Responsável da Inteligência Artificial Generativa na Investigação

Os princípios fundamentais da utilização da IAGen na investigação focam-se na investigação responsável, transparente e de acordo com a lei, baseada nos princípios fundamentais da integridade das atividades de investigação. Além disso, os investigadores devem considerar as seguintes orientações para a utilização de ferramentas de IAGen.

1. Assumir responsabilidade pela produção científica

- Os investigadores são responsáveis pela integridade do conteúdo gerado por/com o apoio de ferramentas de IA, tendo em mente que os sistemas de IA não devem ser autores nem coautores. A autoria implica ação e responsabilidade, pelo que cabe aos investigadores assumir tais princípios. Nesse sentido, terão de ser responsáveis pelo uso apropriado da IA (principalmente como ferramentas de apoio) e pelos resultados que forem partilhados ou publicados.

2. Manter uma visão crítica em relação aos resultados produzidos pela IAGen.

- Os investigadores devem manter uma abordagem crítica relativamente à utilização de resultados produzidos por IAGen, cientes das limitações destas ferramentas, e.g., erros, alucinações e imprecisões.
- Os investigadores devem, também, ter em consideração os principais quadros de valores, padrões culturais e normas sociais incorporadas em cada modelo de IAGen, de acordo com a sua origem e contexto cultural e político.
- É crucial verificar os resultados produzidos pela IAGen, de forma a garantir a exatidão dos mesmos. Tais *outputs* podem ser imprecisos - por exemplo ao citar referências que possam ter sido fabricadas. Em última análise, cabe sempre aos investigadores realizar este processo de verificação.
- Nunca se deve utilizar conteúdo falso gerado por IAGen nos processos científicos,

tais como falsificar, alterar ou manipular dados de investigação originais.

3. Utilizar a IAGen de forma transparente.

- Com vista a garantir a transparência, os investigadores devem detalhar quais as ferramentas de IAGen que foram utilizadas ao longo das atividades de investigação. Por exemplo, utilizar tais soluções como ferramentas básicas de apoio ao autor (e.g., revisão de conteúdo) não deverá ser considerado uma utilização extensiva. No entanto, interpretar análises de dados, realizar revisões de literatura, identificar lacunas na investigação, formular objetivos de investigação, desenvolver hipóteses, etc., podem ter um impacto substancial e, por isso, devem ser explicitamente relatados.
- A referência à ferramenta pode incluir o nome, a versão, a data, etc., bem como informação sobre como foi utilizada e de que forma influenciou o processo de investigação. Caso seja relevante, os investigadores devem disponibilizar o *input* (*prompts*) e o *output*, de acordo com os princípios de *Open Science*.
- Os investigadores devem ter em conta a natureza estocástica (aleatória) das ferramentas de IAGen - ou seja, a tendência para produzir resultados diferentes a partir da mesma informação introduzida. Isto representa um desafio em termos de replicabilidade e robustez dos resultados e conclusões da investigação. Assim, os investigadores devem divulgar ou discutir as limitações das ferramentas utilizadas, incluindo possíveis viés no conteúdo gerado, bem como quaisquer medidas de mitigação.
- A consulta das orientações das revistas científicas onde os investigadores pretendem publicar também é aconselhável, sobretudo no que concerne a utilização de IAGen.

4. Verificar as questões de privacidade, confidencialidade e direitos de propriedade intelectual.

- Os investigadores devem estar cientes de que os dados que introduzem nas ferramentas de IA (texto, dados, *prompts*, imagens, etc.) podem ser utilizados para outros fins, tais como o treino de modelos de IA. Nesse sentido, é importante proteger trabalhos não publicados ou confidenciais, e evitar introduzi-los em sistemas de IA *online* - a não ser que seja garantida a não-utilização ou reutilização desses mesmos dados (por exemplo, para treinar futuros modelos de linguagem, ou para processos de reutilização não-rastreável e não-verificável).

- Os investigadores não devem introduzir dados pessoais de terceiros em sistemas IAGen *online*, a menos que tal seja feito em condições que garantam o cumprimento rigoroso das regras de proteção de dados aplicáveis e dos princípios éticos de autonomia e propriedade.
- Os investigadores devem verificar as opções de privacidade das ferramentas, quem as gere (instituições públicas ou privadas, empresas, etc.), onde são desenvolvidas e as implicações no que toca à introdução de informação nas mesmas. Tal aplica-se a diferentes casos, desde plataformas fechadas, hospedadas em infraestruturas de terceiros com privacidade garantida, até às plataformas abertas, disponíveis *online*.

5. **Cumprir com a legislação nacional, europeia e internacional aquando da utilização de resultados de IAGen.**

- Os resultados gerados por IAGen podem levantar questões relacionadas com a proteção de direitos de propriedade intelectual e os dados pessoais.
- Cabe aos investigadores verificar questões como o plágio (texto, código, imagens, etc.) aquando da utilização de resultados de IAGen. Os *outputs* deste tipo de ferramentas (como os modelos de linguagem de grande escala) podem basear-se em resultados provenientes de outros autores, pelo que poderá ser necessário reconhecer a sua origem e incluir uma citação. Desta forma, os investigadores devem garantir que os resultados utilizados respeitam os direitos de autoria, citando devidamente o trabalho de terceiros.
- Utilizar ferramentas de IAGen para resumir documentos sobre determinado tema poderá ser útil, mas não substitui a consulta de material e fontes legítimas e originais.
- Da mesma forma, os resultados gerados por IAGen podem conter dados pessoais. Caso aconteça, os investigadores são responsáveis por tratar quaisquer dados de forma responsável e adequada, devendo ser cumpridas as regras de proteção de dados em vigor na UE.

6. Evitar a utilização de ferramentas de IAGen de forma extensiva em atividades que possam afetar o trabalho de outros investigadores ou entidades.

- Pedidos para revisão de pares, avaliação de propostas de trabalhos de investigação ou de candidaturas devem ser examinados tendo em conta a capacidade de análise e de julgamento dos próprios investigadores.
- Deve evitar-se a utilização de ferramentas de IAGen para dar resposta a estas tarefas, de modo a eliminar potenciais problemas em termos de tratamento ou avaliação que possam resultar das limitações das mesmas.

7. Aprender, de forma contínua, a melhor forma para utilizar as ferramentas de IAGen.

- As ferramentas de IAGen estão em constante evolução, com novas formas de utilização a ser desenvolvidas. Os investigadores devem manter-se atualizados em termos de aprendizagem e boas práticas, partilhando o seu conhecimento com terceiros ou *stakeholders* relevantes.
- É crucial acompanhar os métodos de desenvolvimento deste tipo de ferramentas, e ter conhecimento sobre os custos energéticos associados, o possível controlo e monopólio por parte de certos agentes, e a legitimidade em termos de recolha de dados.
- Os investigadores e as organizações de investigação devem promover uma cultura de experimentação e aprendizagem em torno da IA, procurando partilhar entre si novas oportunidades.

Concluindo, as ferramentas de IAGen demonstram um grande potencial no que concerne a investigação; no entanto, devem ser utilizadas de forma ponderada, transparente, responsável e criteriosa.

Fontes

INESC TEC

- [Código de Ética do INESC TEC](#)
- [Science & Society: The responsibilities of Responsible AI](#)

Comissão Europeia

- [Guidelines for the Responsible Use of Generative AI in Research – European Research Area Forum \(2024\)](#)

Universidades e Instituições Académicas

- [MIT – Guidelines on AI](#)
- [Universidade Lusófona – IA](#)
- [Harvard University – Guidelines on AI](#)
- [Harvard University – Use of ChatGPT and Other Generative AI Tools](#)
- [Carnegie Mellon University – AI Services and Guidelines](#)
- [KU Leuven – Information on GenAI](#)
- [University of Oxford – Communication Resources and Guidelines on AI](#)