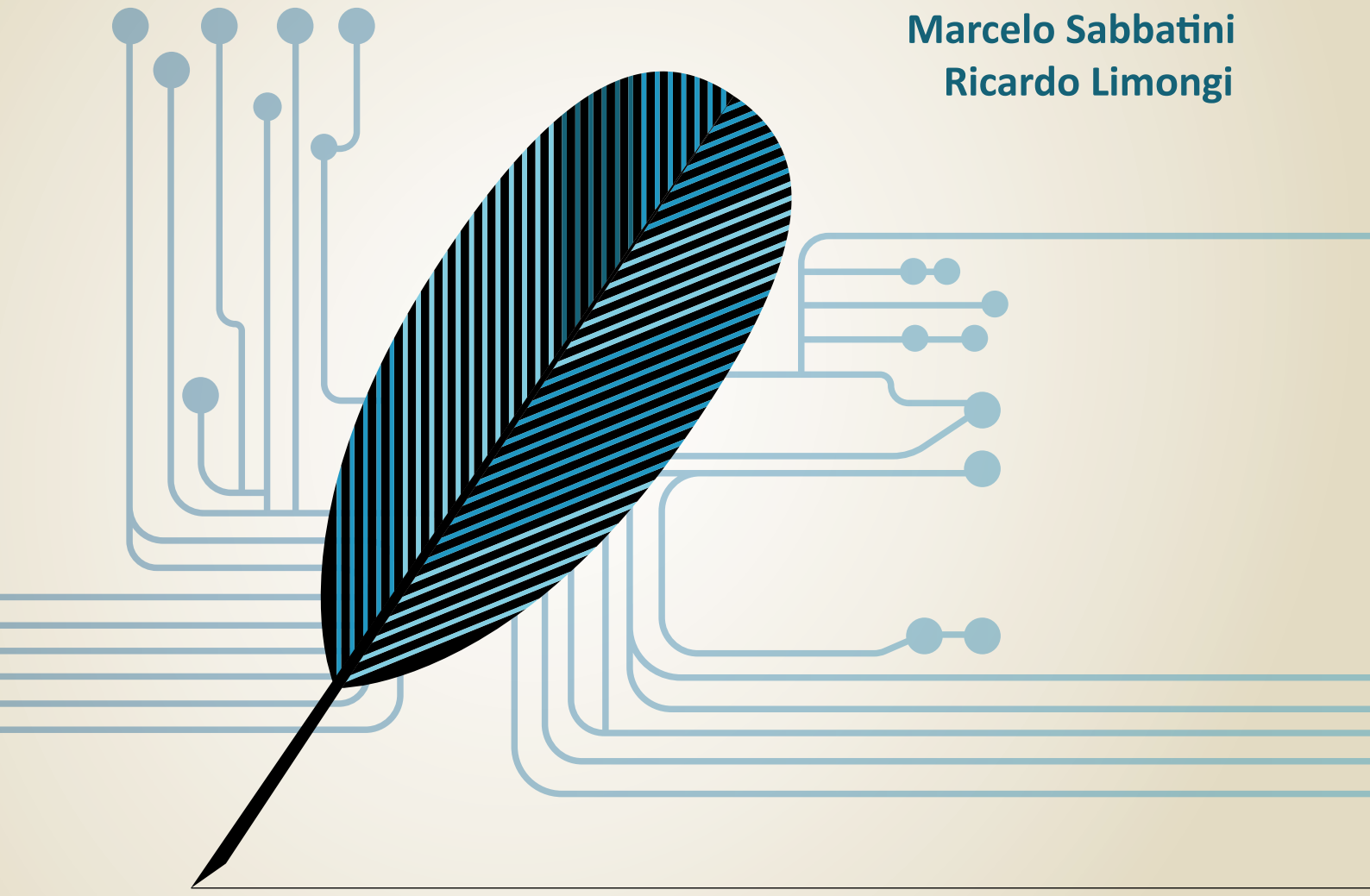
Diretrizes para o uso ético e responsável da Inteligência Artificial Generativa

um guia prático para pesquisadores

Rafael Cardoso Sampaio
Marcelo Sabbatini
Ricardo Limongi



INTERCOM
ABCP Associação Brasileira de Ciência Política
ANPOCS Associação Nacional de Pós-Graduação e Pesquisa em Ciências Sociais

Rafael Cardoso Sampaio
Marcelo Sabbatini
Ricardo Limongi

Diretrizes para o uso ético e responsável da Inteligência Artificial Generativa

um guia prático para pesquisadores

FICHA CATALOGRÁFICA

Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)

Sampaio, Rafael Cardoso
Diretrizes para o uso ético e responsável da inteligência artificial generativa [livro eletrônico] : um guia prático para pesquisadores / Rafael Cardoso Sampaio, Marcelo Sabbatini, Ricardo Limongi. -- São Paulo : Sociedade Brasileira de Estudos Interdisciplinares da Comunicação - Intercom, 2024.
PDF

Bibliografia.
ISBN 978-85-8208-142-6

1. Diretrizes 2. Ética 3. Inteligência artificial - Inovações tecnológicas I. Sabbatini, Marcelo. II. Limongi, Ricardo. III. Título.

24-242426 CDD-006.3

Índices para catálogo sistemático:

1. Inteligência artificial generativa 006.3

Eliete Marques da Silva - Bibliotecária - CRB-8/9380

Esta publicação conta com o apoio da Associação Brasileira de Ciência Política (ABCP) e da Associação Nacional de Pós-Graduação e Pesquisa em Ciências Sociais (Anpocs), que reconhecem a importância da discussão sobre os impactos da inteligência artificial sobre as pesquisas acadêmicas e que há a necessidade de maturação do debate na temática, notadamente em termos de boas práticas e diretrizes, como pretende o documento em questão.

SUGESTÃO PARA CITAÇÃO

SAMPAIO, R.C.; SABBATINI, M.; LIMONGI, R. **Diretrizes para o uso ético e responsável da Inteligência Artificial Generativa:** um guia prático para pesquisadores. São Paulo: Editora Intercom, 2024.

Rafael Cardoso Sampaio
Marcelo Sabbatini
Ricardo Limongi

Diretrizes para o uso ético e responsável da Inteligência Artificial Generativa

um guia prático para pesquisadores

Sumário

Agradecimentos

Agradecemos à diretoria do Intercom, na figura de seu presidente Juliano Domingues, pela acolhida do material e pelo esforço de Nara Cabral Scabin para a publicação em curto período. Da mesma forma, agradecemos o apoio da diretoria da Associação Brasileira de Ciência Política (ABCP), através de seu presidente, Bruno Reis, e da diretoria da Associação Nacional de Pós-Graduação e Pesquisa em Ciências Sociais (Anpocs), através de seu presidente Adriano Codato.

Agradecemos aos vários colegas do grupo “Ferramentas de IA para pesquisa” pelas inúmeras dicas e sugestões e, em especial, à Angelita Alice Jaeger pela generosa leitura e revisão e a Daniel Fugisawa pelas inúmeras dicas e indicações. Também agradecemos aos integrantes do grupo de pesquisa Margem da UFMG por suas críticas e sugestões valiosas, notadamente nas figuras de Maria Alejandra Nicolás, Ricardo Fabrino e Victoria Lima. Damos um agradecimento especial a Cristiane Sinimbu, que gentilmente leu, revisou e normatizou o texto em um curto espaço de tempo.

Agradecemos ainda aos diversos pesquisadores, professores e alunos pelas interações significativas ao longo do último ano em palestras, mesas-redondas, eventos acadêmicos e redes sociais. Parafraseando Paulo Freire, não podemos temer o debate; é numa discussão criadora de análise da realidade que precisamos enfrentar os desafios impostos pela Inteligência Artificial Generativa.

Prefácio

Uma das minhas primeiras memórias sobre pesquisa acadêmica tem a ver com tecnologia. No início dos anos 2000, um tio que fazia mestrado me mostrou uma de suas atividades: entre animado e preocupado, ele exibia as centenas de artigos científicos que encontrava no Google, dizendo que precisava lê-los um a um. Ao percorrer as infindáveis páginas de textos sobre solos, cultura de milho e processos de lixiviação encontrados no recém-criado buscador, ele argumentava sobre a importância da "ferramenta" tecnológica para que a ciência pudesse avançar e para que pessoas como ele pudessem ter acesso ao que se produz no mundo todo.

Para pessoas como eu, que aprenderam a digitar "no computador" enquanto aluno de graduação, aquele tipo de transformação adicionava novas camadas potencialmente revolucionárias às mudanças em curso. Escrever teclando, em um suporte vertical, flexível e não linear já me mostrara como nossos jeitos de pensar poderiam ser profundamente afetados por tecnologias. Ao estudar história social dos meios de comunicação, tive a chance de pensar mais criticamente sobre os modos como tecnologias da inteligência são parte muito importante de profundas mudanças sociais e históricas, desde a escrita à fotografia, passando pela televisão e as mídias digitais. As formas de produção, circulação e consumo do conhecimento não são nada estanques e implicam alterações significativas em padrões de subjetividade, formas de socialização, estruturas institucionais e bases epistêmicas que alicerçam não apenas escolhas, mas horizontes de possíveis.

À época, seria difícil prever a velocidade e profundidade com que tanta coisa mudaria em duas décadas. Entre o aprendizado da escrita não linear, o primeiro contato com um "repositório" de artigos de periódicos e a popularização dos LLMs, houve muitos desdobramentos tecnológicos que deixaram marcas sobre práticas, processos, aspirações, anseios e riscos a atravessar a vida acadêmica. Para citar alguns exemplos desses desdobramentos, pense-se na digitalização de enormes volumes de dados e na *datificação* de tantas experiências, na plataformização de comunidades diversas, na *smartphonização* de tantas atividades, no uso generalizado de sistemas automatizados, na quantidade de reuniões mediadas por plataformas diversas, nos aplicativos de gestão de trabalho coletivo, ou mesmo no declínio das outrora essenciais "pastas de xérox" das universidades. Tudo isso afeta a forma como delineamos ações de ensino, pesquisa, extensão e administração nas universidades.

É nesse cenário que uma onda de incerteza parece se abater de forma ampla sobre institutos de pesquisa, universidades e revistas acadêmicas atualmente. O avanço da Inteligência Artificial Generativa e sua centralidade no debate público colocaram várias questões às instituições acadêmicas mundo afora. Frequentemente explicitado pelo temor generalizado de que "estudantes" entreguem trabalhos de que não são autores, o impacto da IA Generativa sobre Ensino, Pesquisa, Extensão e Administração Acadêmica é muito mais profundo e inclui dilemas ambientais, questões de direitos autorais, proteção de dados e os perigos de aninhar camadas de opacidade no coração de processos não inteiramente controláveis por pesquisadoras e pesquisadores. Modelos de IA podem trazer velocidade, eficiência e viabilizar atividades que nem se imaginava serem possíveis, mas também afetam nossas compreensões sobre autoria, validação, replicação e sistematicidade. Eles podem aumentar a precisão de certos processos, mas também inserir alucinações e vieses não perceptíveis ao longo de diversas ações.

Esse contexto torna imperativo que instituições acadêmicas enfrentem a difícil tarefa de debater como lidarão com a IA e seus impactos em suas múltiplas atividades. Justamente por isso, o livro de Rafael Sampaio, Marcelo Sabbatini e Ricardo Limongi chega em bora hora. Ao refletir sobre Diretrizes para uso ético e responsável de IA Generativa, eles se debruçam sobre desenvolvimentos tecnológicos contemporâneos e abordam alguns dos dilemas que atravessam o cotidiano de professores e pesquisadores. Como bem alertam os autores, o livro não pretende se converter em manual estanque ou encerrar o debate sobre as questões que aborda. Ao contrário, ele é um convite de abertura à discussão, que interpela instituições (como o MEC, a Capes, o CNPq) e pesquisadores a encarar, de forma franca e responsável, a fluidez das bases sobre as quais assenta seu fazer.

O convite é para uma reflexão coletiva ampla e continuada. Na UFMG, instituição à qual estou vinculado, uma Comissão tem trabalhado, desde novembro de 2023, na construção de uma Política de Inteligência Artificial. Entre as atribuições da Comissão, estão a articulação de pesquisas, a indução de projetos de extensão e a estruturação dinâmica de diretrizes. Além disso, a Comissão tem proposto formas de escutar a comunidade acadêmica para abordar dilemas e percepções sobre IA em diferentes áreas do conhecimento. Defende-se, ainda, a articulação de uma rede de universidades que faça esse processo em conjunto. A obra de Sampaio, Sabbatini e Limongi pode ser pensada como peça relevante desse quebra-cabeças sobre o qual precisaremos nos debruçar coletivamente. Ela instiga o debate e nos convoca a assumir uma tarefa incontornável para pensar Ensino, Pesquisa, Extensão e Administração Acadêmica hoje. Uma tarefa de que dependem os próprios rumos das universidades.

Ricardo Fabrino Mendonça
Professor de Ciência Política
Presidente da Comissão Permanente de Inteligência Artificial da UFMG

Apresentação

Diante dos desafios e preocupações que o uso da Inteligência Artificial Generativa (IAG) tem colocado para a comunidade científica, em termos de integridade acadêmica, originalidade e autoria, privacidade, entre outros, visamos orientar pesquisadores de todos os campos do conhecimento sobre o uso ético e responsável dessas tecnologias.

Dessa forma, partimos de uma apresentação dos princípios gerais do que consistiria a ética e responsabilidade na pesquisa científica, a partir do uso da IAG. E logo, num segundo momento, oferecemos diretrizes práticas para pesquisadores que utilizam ferramentas de IAG em suas pesquisas, principalmente, para o fortalecimento de uma cultura que evidencie o foco no processo científico e a IAG como meio para pesquisas com melhores contribuições. Além de desmistificar o entusiasmo exagerado (*hype*) em torno da IAG, buscamos contribuir para o uso soberano da IA no Brasil,

Para atingir estes objetivos, o guia está dividido em diferentes seções. Na "Introdução", contextualizamos a rápida ascensão da IAG, a exemplo do ChatGPT, e avaliamos brevemente seus impactos na pesquisa científica. Em "Princípios Gerais", discutimos aqueles que consideramos os pilares para seu uso ético e responsável, a saber: compreensão das ferramentas de IAG, autoria humana, transparência, integridade da pesquisa acadêmica, plágio, originalidade, direitos autorais, preservação da agência humana, uso eticamente orientado e letramento em IA.

Entre os temas abordados, destacamos o risco dos vieses e alucinações, inclusive com a possibilidade de fabricação de dados, além dos cuidados com a privacidade e a segurança, tópicos essenciais em se tratando de dados sensíveis e resultados inéditos. A discussão sobre a integridade acadêmica também é de grande relevância, estando relacionada com a autoria e a originalidade das pesquisas. Somando a compreensão profunda das ferramentas, com seus limites e implicações éticas, alertamos para o risco da dependência tecnológica, com a inibição de habilidades críticas e do pensamento independente. Em conjunto, estes princípios atendem a uma necessidade mais ampla de validade, transparência e replicabilidade na pesquisa científica, o que será alcançado somente com a preservação da Agência Humana e com o uso da tecnologia como ferramenta complementar, nunca em substituição ao pesquisador, o que enfatiza a necessidade de letramento nos sistemas de IAG.

Logo, em "Princípios Práticos" apresentamos orientações para diferentes usos da IAG na pesquisa científica, com atenção para os cuidados a serem tomados. Neste sentido, destacamos o uso para a exploração inicial de ideias (*brainstorming*), busca de materiais acadêmicos - assim como sua leitura e resumo - seguida da escrita acadêmica, análise e apresentação de resultados, programação, transcrição, tradução, elaboração de pareceres e avaliações, uso na forma de agentes de IA e, finalmente, detecção de IA Generativa.

A título de "Considerações finais", analisamos as atuais políticas acadêmicas para o uso da IAG no contexto científico e propomos uma abordagem colaborativa para o uso ético e responsável dessas ferramentas. E como "Palavras Finais", discutimos os desafios do *hype* em torno da IA e a importância do uso soberano de IAG no Brasil. Neste ponto, destacamos a necessidade de investimentos em recursos humanos e infraestrutura para o desenvolvimento da IA no Brasil e convidamos pesquisadores, universidades, associações científicas e órgãos de fomento à pesquisa para que se engajem em discussões e ações conjuntas para garantir o uso ético e responsável.

A construção de uma IA nacional, livre de vieses e com respeito à soberania do país, é fundamental para que a pesquisa científica brasileira se beneficie de forma justa e segura das inovações da IAG.

Preâmbulo

Lendo um trabalho enviado a seu grupo de trabalho num evento, a atenção do avaliador é chamada por uma citação particularmente interessante. Ele vai até as referências bibliográficas, copia o nome dos autores e título e busca na Internet. Mas não recebe nenhuma correspondência exata. Pesquisa pelo nome da revista, o resultado é o mesmo, ela parece não existir. Desconfiado, repete o processo para as outras referências, nenhuma delas era verificável.

(...)

Após a longa espera, finalmente, o pesquisador recebe os pareceres do artigo que enviara à revista. Porém, uma frase estranha destoa, em meios às observações e críticas. "Como modelo de Inteligência Artificial…". Afinal, quem escreveu o parecer?

(...)

O artigo de pesquisa foi publicado numa revista importante, de uma grande editora. Mas diferentemente da maioria dos textos ultra especializados, de audiência seleta, alcançou milhões de visualizações. O motivo? Uma ilustração cientificamente inexata, com toques surreais, acompanhada de texto indecifrável.

(...)

O texto do pré-projeto de pesquisa enviado ao processo seletivo da pós-graduação aparentava ter boa textualização, sem erros gramaticais e ortográficos. Porém, notava-se uma falta de coesão entre as ideias, as questões teórico-metodológicas eram tratadas de forma superficial. Junto a uma ausência de exemplificação, a leitura do texto também chamava a atenção pelo excesso de adjetivos e advérbios. Especialmente um que se repetia: "crucial".

(...)

Um pesquisador está utilizando um agente conversacional para ajudar a estruturar a apresentação que fará num congresso. Porém, ele se surpreende ao ver que a IA já conhece sua proposta, uma tese inovadora que ele não publicou em lugar nenhum. Quer dizer, já havia enviado o artigo para avaliação numa revista, mas fora isso ninguém havia lido o texto.

(...)

Após o escândalo do plágio revelado justamente no momento da defesa de doutorado, a investigação aprofundada descobriu o que havia acontecido. O candidato havia utilizado geração automática de texto. Porém, não percebera que aquelas ideias muito específicas "tinham dono". Mas de alguma forma, a Inteligência Artificial havia aprendido.

(...)

Longe de serem ficção, os casos relatados representam o cotidiano da pesquisa científica, na era pós-GPT. Tendo em comum o uso da Inteligência Artificial Generativa de forma descuidada, sem o conhecimento embasado de como a tecnologia funciona, de quais situações ela mais se presta ao uso e de quais deveriam ser evitadas, as consequências no plano da ética acadêmica são devastadoras.

Introdução

Um meteoro, capaz de impactar todos os aspectos da vida social, incluindo as formas de produção e circulação do conhecimento científico? O lançamento público do ChatGPT em 30 de novembro de 2022 foi amplamente percebido dessa forma. Segundo Santaella, "a metáfora não é sem razão. No decorrer desse mês e naqueles que se seguiram, foi desmedido o volume de notícias, de colunas jornalísticas, de blogs e dos primeiros textos de natureza interpretativa" (Santaella, 2023, p. 9). E ela afirma que "até hoje o impacto não cessou, ao contrário, [...] tornando-se inclusive apenas a ponta do iceberg de uma grande quantidade de sistemas em competição, além de outros subsidiários" (Santaella, 2023, p. 9).

Como tecnologia digital de mais rápida adoção na história, o agente conversacional baseado em Inteligência Artificial bateu a marca de 1 milhão de usuários em apenas cinco dias[1], atingindo 100 milhões de usuários em apenas dois meses e superando outros gigantes da tecnologia, como Instagram e TikTok[2]. Diferentemente de chatbots repetitivos e assistentes pessoais automatizados e limitados, o ChatGPT, em especial, mostrou ser capaz de responder e gerar diálogos semelhantes aos humanos, mantendo conversas prolongadas.

Além disso, esse e outros modelos de IA generativa, a exemplo de *Claude*, *Copilot*, *Gemini*, *Llama* e *Maritalk*, possibilitam a automação de diversas atividades, como a redação de mensagens pessoais, preenchimento de formulários, criação de textos padrão, tradução, resumo, síntese, organização e estruturação de conteúdos, análise extensiva de dados, bem como a criação e correção de códigos de programação – tudo isso por meio de comandos em linguagem natural, chamados *prompts* (Dwivedi *et al.*, 2023; Ramos, 2023; Santaella, 2023; Sohail *et al*., 2023; Susarla *et al.*, 2023; Almeida, Nas, 2024; Limongi, 2024; Sampaio *et al*., 2024a).

Porém, tais possibilidades de uso também suscitaram quase imediatamente uma reação de desconfiança e temor. Especialmente os conceitos de autoria, originalidade e integridade acadêmica foram abalados, com pesquisadores e professores questionando quais seriam os efeitos e consequências da automação dos produtos e subprodutos da pesquisa científica. A discussão se acirra, conforme a IAG afeta aspectos que não esperávamos serem realizados por máquinas, como a criatividade na escrita e em atividades artísticas (Gonçalves, 2023; Santaella, 2023; Habib *et al*., 2024).

Ao analisar as diretrizes de universidades americanas, Soares *et al*. (2023) destacam como algumas delas reconhecem essas ferramentas como facilitadoras do processo de pesquisa, especialmente como geradoras de resumos que simplificam a descoberta de novos documentos e a investigação de fontes relevantes que podem adensar o conteúdo e a pesquisa, podendo inclusive ser usadas para "a geração de ideias bem como para a estruturação de conteúdo, utilizando ferramentas de IA para desenvolver e organizar ideias de forma mais coesa" (Soares *et al*., 2023, p. 83).

Então, qual é exatamente o problema do uso da IAG em práticas acadêmicas? Ora, pesquisadores já sabem há bastante tempo que o mundo contemporâneo é fortemente influenciado por algoritmos, frequentemente baseados em aprendizagem de máquina (*machine learning*) e Inteligência Artificial Preditiva, modalidade voltada para fazer previsões e decisões baseadas em dados[3]. Seu uso é bastante presente em redes sociais digitais e em aplicativos utilizados na vida cotidiana, a exemplo de Waze, Google Maps e Uber. Na esfera do consumo cultural, a tecnologia também está presente em mecanismos de recomendação de YouTube, Netflix e Spotify, conhecidos do grande público (Lemos, 2021; Mendonça, Filgueiras, Almeida, 2023). O que mudou então?

1. NOGUEIRA, João Gabriel. ChatGPT bateu 1 milhão de usuários em apenas 5 dias, diz OpenAI. **TecMundo**, 06 dez. 2022. Disponível em: https://www.tecmundo.com.br/software/259681-chatgpt-bateu-1-milhao-usuarios-5-dias-diz-openai.htm Acesso em: 30 set. 2024.

2. FORBES TECH. ChatGPT tem recorde de crescimento da base de usuários. **Forbes Brasil**, 02 fev. 2023. Disponível em: https://forbes.com.br/forbes-tech/2023/02/chatgpt-tem-recorde-de-crescimento-da-base-de-usuarios/. Acesso em: 30 set. 2024.

3. Arvind Narayanan e Sayash Kapoor (2024), pesquisadores da Universidade de Princeton e vozes influentes no debate da IA, argumentam que a IA Preditiva é "óleo de cobra", um remédio que promete curas milagrosas, mas não funciona como o anunciado. No caso, a razão do fracasso seria a dificuldade inerente de se prever o comportamento social humano; este não deveria ser um problema tecnológico, em primeiro lugar, ainda mais considerando a forma como pode afetar a vida das pessoas. Ainda que seja usada na pesquisa científica, inclusive de forma temerária como alertam estes autores, a IA Preditiva não será objeto deste guia.

Antes do lançamento do ChatGPT, apenas programadores e desenvolvedores de plataformas trabalhavam diretamente com tais tecnologias e eram capazes de verificar os potenciais da IA Generativa e Preditiva; logo o cidadão comum não tinha qualquer noção de tais sistemas. O ChatGPT, "por outro lado, caiu direto no colo de qualquer ser humano [...] [que] pode ver tarefas, que lhe custavam tempo, sendo realizadas rapidamente por um sistema solícito e prestativo" (Santaella, 2023, p. 22), ainda mais se tratando de um sistema gratuito, e de fácil uso.

Esses Grandes Modelos de Linguagem (*Large Language Models* ou LLMs, no original em inglês) basicamente usam o seu treinamento para reconhecer a entrada fornecida pelo usuário (normalmente, chamada de *prompt*), avaliando quais são as palavras mais importantes do pedido original e, com base no seu treinamento, a resposta mais parecida com a de um ser humano, ao menos no sentido estatístico. Esses sistemas implementam modelos generativos, que são coleções de princípios e diretrizes que adquirem conhecimento sobre a coocorrência convencional de palavras. Esses princípios são empregados para antecipar a formação de novos textos, procurando replicar com a maior precisão possível a maneira pela qual a linguagem humana é utilizada[4]. Para facilitar essas capacidades preditivas, os modelos passam por um treinamento extensivo utilizando gigantescos conjuntos de materiais, geralmente com bilhões, senão trilhões de parâmetros (*tokens*)[5], o que os permite ser muito mais sofisticados e expressivos que modelos anteriores (Almeida *et al*., 2023; Liao, Vaughan, 2023; Ray, 2023; Sohail *et al*., 2023; Khan *et al*., 2024).

O ChatGPT e outros grandes modelos de linguagem, como Claude, Copilot, Gemini e Llama, Maritalk, apresentam semelhanças com um "papagaio estocástico" que se esforça para replicar informações adquiridas por meio de treinamento extensivo, mas que efetivamente não compreende o que está dizendo (Ray, 2023; Sohail *et al*., 2023; Dabis, Csáki, 2024)[6]. O modelo constrói respostas de forma incremental, utilizando uma estrutura de probabilidades para encontrar uma sequência de signos que pareça fazer sentido para um observador humano. Por esse motivo preciso, esses modelos raramente reproduzem as respostas literalmente e, às vezes, cometem erros em suas respostas, resultando no que vem sendo denominado "confabulações", ou mais popularmente, "alucinações". Essas ocorrências ocorrem quando o algoritmo seleciona uma probabilidade sintaticamente válida, mas factualmente imprecisa (Alkaisse, Mcfarlane, 2023; Liao, Vaughan, 2023; Santaella, 2023).

Aplicadas ao nosso contexto acadêmico, ferramentas de inteligência artificial são aquelas que ajudam o pesquisador a realizar diferentes partes da pesquisa científica, ou ainda, instrumentos de apoio para facilitar ou acelerar certos trabalhos, emulando a função de assistentes de pesquisa. A expectativa é que sejam assistentes e não supervisores da pesquisa.

Entretanto, não podemos apenas aceitar uma visão de que tais IAs seriam apenas ferramentas, a exemplo de uma calculadora ou um computador. Como já dito, elas exibem comportamentos inteligentes, podendo nos prover textos, respostas, imagens, análises e sugestões de todos os tipos. Essa mímica de um comportamento inteligente é fortemente baseada na forma como essas ferramentas foram construídas e treinadas. Em outras palavras, tais ferramentas exibem um comportamento inteligente porque elas são ensinadas com grandes volumes de dados a tomar decisões parecidas com humanos em decisões similares, ainda que dependam de forte poder computacional em termos de programação e mesmo de *hardware* (Unesco, 2024; União Europeia, 2024; Sohail *et al*., 2023; Susarla *et al*., 2023).

Isso traz quatro questões importantes a serem consideradas, que são base deste guia. A primeira é que a maior parte dos Grandes Modelos de Linguagem são proprietários, pertencendo às Big *Techs*, grandes empresas de tecnologia com sede no Vale do Silício nos Estados Unidos (Morozov,

4. Neste repositório do Github, estão reunidos os principais *papers* a explicarem técnicas de aprendizado de máquina, LLMs e similares. Disponível em: https://github.com/dair-ai/ML-Papers-Explained. Acesso em: 8 out. 2024.

5. Durante a escrita deste guia, já se nota que o mercado de modelo de linguagens também começa a investir em pequenos modelos de linguagem, pensados para tarefas mais simples, rápidas e para rodarem em aparelhos com menor capacidade processional, a exemplo de celulares. Dito de outra forma, não será necessário acessar a tecnologia de IA por meio da Internet, ela estará "embutida" em nossos dispositivos. Confira: THOMAS, Rosemary J. Smalll but powerful: a deep dive into Small Language Models (SLMs). **Medium**, 06 dez. 2023. Disponível em: https://medium.com/version-1/small-but-powerful-a-deep-dive-into-small-language-models-slms-b793bdb002f2. Acesso em: 8 out. 2024.

6. Para discussões mais aprofundadas sobre como as IAGs podem exibir padrões de raciocínio diferente dos humanos, ver Kauffman (2022); Röhe, Santaella (2023) e Santaella (2023).

2018). Com efeito, isso significa que a maior parte das ferramentas de IA tenderão a privilegiar visões de mundo estadunidenses, ou pelo menos dos países hegemônicos ocidentais. Para a academia, isso tem impactos de privilegiar a visão científica desses centros de pesquisa, frequentemente apresentando uma lógica quantitativista, baseada em experimentos ou quasi-experimentos. Isso pode trazer dificuldades a certos grupos, como pesquisadores qualitativos e/ou das Ciências Humanas tradicionais, só para citar alguns (Sampaio *et al*., 2024b; Unesco, 2024).

A segunda questão é que as *Big Techs*, para oferecer esse serviço "gratuitamente" ou a preços acessíveis, aproveitam as interações com humanos para treinamento de seus modelos. Isto significa que, em maior ou menor medida, os dados disponibilizados e as interações com pesquisadores se tornarão parte do sistema. Essa preocupação é notoriamente mais complicada quando estamos falando de pesquisas de ponta, com dados inéditos, alimentando a lógica de um "colonialismo de dados" (Cassino, Souza, Silveira, 2021; Oliveira, Neves, 2023). Além disso, há uma preocupação com pesquisas que contêm dados sensíveis, a exemplo de certos tipos de entrevista em profundidade ou grupos focais (Resnik, Hosseini, 2024; Sampaio *et al*., 2024b; Unesco, 2024; União Europeia, 2024), e até mesmo estudos científicos em desenvolvimento.

A terceira questão a ser observada é o viés (*bias*) presente nos bancos de dados e nos algoritmos. Em outras palavras, se os dados primordiais para o treinamento de tais modelos vieram da Internet, então isso significa que problemas gerais da sociedade serão refletidos na forma como esses modelos irão reagir. Dito de outra forma, os grupos subrepresentados, excluídos e marginalizados, a exemplo de mulheres, negros, povos indígenas, LGBTQIAPN+, entre outros, que não possuem representação adequada na Internet, também passarão por uma nova marginalização nos Grandes Modelos de Linguagem. Ao contrário do *hype*[7] geralmente propagado de serem uma "tecnologia disruptiva", na prática, as IAs Generativas tendem a agir para manter o *status quo*, reforçar dinâmicas e estruturas de poder existentes e agravar o problema para tais grupos (Dwivedi *et al*., 2023, Kaufman, 2022; Silva, 2022; Unesco, 2024).

A quarta questão é que tais modelos geralmente dão uma resposta ao pedido (*prompt*) original, mesmo que não haja dados suficientes para um retorno verdadeiro. Ou seja, tais modelos podem errar ou "mentir", inventando fatos e mesmo referências bibliográficas inexistentes! Denominado de "confabulação" ou "alucinação"[8], o fenômeno ocorre por uma raridade estatística, quando a IA tenta imitar um humano, sem ter treinamento suficiente, e por terem como foco a performance, velocidade da resposta em detrimento da melhora na taxa de acerto, inventando respostas (Alkaissi, McFarlane, 2023; Ray, 2023; Sohail *et al*., 2023; Lemos, 2024; União Europeia, 2024).

Neste ponto, cabe uma consideração: a IA Generativa tem como característica a capacidade de

> transformar textos em forma de narrativa, extraindo-os de fontes possivelmente confiáveis, usando fatos possivelmente verídicos, aplicando teorias possivelmente verdadeiras e visualizando-os através das lentes de uma possível análise crítica (Cope, Kalantzis, 2023, p. 843, tradução nossa).

Como a citação direta trata de preservar com a repetição do termo "possível", um texto generativo é, sobretudo, resultado de um cálculo probabilístico. Assim, em maior ou menor medida, atende a uma realidade socialmente estabelecida, mas não necessariamente representa uma verdade empírica.

Considerando esses quatro aspectos e o potencial da IA Generativa para diversas etapas do fazer científico, muitos acadêmicos vêm buscando testar tais ferramentas e verificar as suas habilidades e utilidades para o cotidiano científico (Cong-Lem *et al*., 2024; Rodrigues *et al.*, 2024; Sampaio *et al*., 2024b; Hassan *et al*., 2024; Khan *et al*., 2024). Tal experimentação, entretanto, vem acompanhada por uma discussão presente em boa parte da ciência a respeito de seu uso ético e responsável (Franco *et al*., 2023; Soares *et al*., 2023; Santaella, 2023; Resnik, Hosseini, 2024; Stahl, Eke, 2023; Chetwynd, 2024), incluindo a preocupação com a produção acelerada de materiais acadêmicos de baixa qualidade ou mesmo plagiada apenas para pontuação de currículos, alimentando ainda mais o fenômeno do produtivismo acadêmico (Cotton *et al*., 2023; Curtis, 2023; Velez, Rister, 2024).

Diversos periódicos, associações científicas, editoras, universidades e mesmo órgãos multilaterais, como

7. Significa entusiasmo exagerado, o termo é uma concentração de hipérbole, figura de linguagem utilizada para denotar ênfase e exagero.

8. André Lemos (2024) busca complexificar a questão, evidenciando que podemos definir em "erros, falhas e perturbações digitais", que frequentemente são inerentes aos objetos. Aqui, optamos pela simplificação e usaremos alucinação ou confabulações.

a ONU e União Europeia, já produziram materiais estabelecendo princípios e recomendações de uso (Moorhouse *et al.*, 2023; Soares *et al*., 2023; Dabis, Csáki, 2024; Perkins, Roe, 2024; Limongi, 2024; Unesco, 2024; União Europeia, 2024; Velez, Rister, 2024). Por iniciativa própria, algumas universidades[9], a exemplo de PUC-SP[10], SENAI Cimatec (Bahia)[11], Universidade Federal de Minas Gerais (UFMG)[12] e Universidade de São Paulo (USP)[13] apresentaram suas próprias normativas e recomendações, sob o consenso de que professores e pesquisadores se tornem agentes da disseminação de boas práticas para alunos e jovens pesquisadores (Alves, 2023; Fonseca e Campiglia, 2023; Unesco, 2023; União Europeia, 2024).

Afinal, "proibir deve estar entre os piores caminhos, pois acaba por incentivar o uso oculto", enquanto "ignorar significa alienar-se de um problema que exige atenção e encaminhamentos". Dessa forma, a ética surge como elemento principal e obrigatório, exigindo, "antes de tudo, informar-se, conhecer, experimentar e avaliar para melhor agir" (Santaella, 2023, p. 22).

Em outras palavras, não faz qualquer sentido banir a tecnologia para estudantes e jovens pesquisadores que vivem em um mundo em que a tecnologia é ubíqua e tem papel fundamental (Dwivedi *et al*., 2023). Contudo, tampouco, significa que devamos abraçá-las apenas por serem "novas" ou "inovadoras" sem uma perspectiva ética e crítica[14]. Parece-nos que há uma contradição na qual as IAG são passíveis de serem objetos de pesquisa, porém não de serem incorporadas nas práticas acadêmicas. Em eventos e outras interações cotidianas, foram frequentes os relatos de pesquisadores sobre um sentimento de "policiamento" por parte dos colegas, que muitas vezes também usam as ferramentas em segredo (Velez, Rister, 2024).

A proibição do ChatGPT e de outros modelos generativos é contrastado com implicações mais profundas na mudança de concepção da natureza da produção do conhecimento científico. Neste ponto, outro levantamento bibliográfico identificou a necessidade do estabelecimento de diretrizes éticas, a colaboração interdisciplinar e a promoção de uma educação crítica, reflexiva e consciente sobre o uso de IAG, numa perspectiva de letramento. Estas seriam as condições para estabelecer uma pesquisa responsável, que vá de encontro aos valores da justiça social e integridade acadêmica. (Souza, Sabbatini, 2024).

Nesse sentido, a formação de comitês éticos multidisciplinares para supervisionar especificamente o desenvolvimento e a aplicação da IA na pesquisa por universidades, instituições de pesquisa e órgãos financiadores é uma necessidade premente, ainda que escassamente realizada. Um levantamento sobre diretrizes presentes nas universidades mais bem ranqueadas do mundo mostrou que mais da metade delas já tinham políticas de IAG, como uma resposta para manter a integridade acadêmica e adaptar as práticas avaliativas a um cenário tecnológico em rápida evolução (Moorhouse *et al.*, 2023). O estudo mostra que diretrizes bem estruturadas ajudam professores a projetar avaliações mais resistentes ao uso indevido de IAG e incentivam a transparência dos alunos em seu uso e recomenda que as avaliações integrem a IA de forma controlada, promovendo tanto a ética acadêmica quanto o letramento digital de alunos e pesquisadores em formação.

Assim, é identificada a necessidade de se constituir um *ethos* da sabedoria prática (*phornesis*) que, em última instância, tenha como resultado uma formação ética e de integridade acadêmica. Aludindo ao sentido de prudência e sensatez, na

9. A Revista Pesquisa Fapesp produziu uma interessante reportagem sobre o assunto. Confira em: SCHMIDT, Sarah. Universidades brasileiras discutem regras de uso de inteligência artificial. **Pesquisa FAPESP**, 11 set. 2024. Disponível em: https://revistapesquisa.fapesp.br/universidades-brasileiras-discutem-regras-de-uso-de-inteligencia-artificial/. Acesso em: 2 out. 2024.

10. PUC-SP. Manual Ético para o uso da Inteligência Artificial Generativa. **TECCOGS - Revista Digital de Tecnologias Cognitiva**s, São Paulo, n. 28, 2023. Disponível em: https://revistas.pucsp.br/index.php/teccogs/issue/view/2973/495. Acesso em: 2 out. 2024.

11. SENAI CIMATEC. **Guia para uso de IA Generativa no Centro Universitário SENAI CIMATEC**. Salvador, BA. Disponível em: https://seja.senaicimatec.com.br/wp-content/uploads/2024/03/GUIA-DE-IA-NA-EDUCACAO.pdf. Acesso em: 2 out. 2024.

12. Universidade Federal de Minas Gerais (UFMG). **Recomendações para o Uso de Ferramentas de Inteligência Artificial nas Atividades Acadêmicas na UFMG**. Belo Horizonte, MG. 2024. Disponível em: https://wwwufmg-hml.dti.ufmg.br/ia/wp-content/uploads/2024/09/Uso-de-Ferramentas-de-IA-na-UFMG.pdf. Acesso em: 2 out. 2024.

13. A USP tem se dedicado a debater o tema em inúmeros seminários, mesas-redondas e afins, notadamente por meio de seu Instituto de Estudos Avançados, conforme o dossiê presente nesta matéria de PRADO, Luiz. Dossiê antecipa as possibilidades e os riscos da inteligência artificial. **Jornal da USP**, 7 jun. 2024. Disponível em: https://jornal.usp.br/cultura/dossie-antecipa-as-possibilidades-e-os-riscos-da-inteligencia-artificial/. Acesso em: 14 out. 2024. Em especial, podemos destacar a cátedra de inteligência artificial responsável, liderada por Virgílio Almeida, que tem produzido fortemente no tema: https://iaresponsavel.iea.usp.br/producao/. Acesso em: 14 out. 2024.

14. Recomendamos a leitura de uma matéria pertinente sobre essa questão. OLIVEIRA, Ruam. Sem diretrizes, ensino superior tenta entender impacto da inteligência artificial. **Porvir: Inovações em Educação**, 23 jun. 2023. Disponível em: https://porvir.org/sem-diretrizes-ensino-superior-ainda-tenta-entender-impacto-de-inteligencia-artificial/. Acesso em: 2 out. 2024.

tradição filosófica da obra clássica de Aristóteles, a proposta é evitar uma abordagem legalista e cultivar a "dedicação, esforço, compromisso, responsabilidade, conduta, partilhas de experiência e constância nos hábitos" que resultem num princípio formativo que seja "bom e virtuoso" (Tedesco, Ferreira, 2023, p. 21-22).

De forma similar, após extensa análise de diretrizes existentes, Franco, Viegas, Röhe *et al*. (2023) propõem um conjunto abrangente de recomendações para a implementação e governança de Inteligência Artificial Generativa em instituições acadêmicas (ver também Nas, Almeida, 2024). Os autores enfatizam a necessidade de políticas institucionais claras, alinhadas com a missão organizacional e compatíveis com normas regulatórias. Destacam a importância de medidas proativas de segurança de dados e conformidade legal, especialmente com legislações como a Lei Geral de Proteção de Dados no Brasil (LGPD).

Estas diretrizes estão alinhadas com uma pesquisa desenvolvida pela Oxford University Press (2024) focada em captar as atitudes e usos da IAG entre pesquisadores de diferentes disciplinas e estágios de carreira. Realizada entre março e abril de 2024, o levantamento utilizou um questionário online, com entrevistas cognitivas para garantir clareza e precisão nas perguntas (n=2.345). O estudo revelou perfis variados de pesquisadores em relação à adoção e ao ceticismo em torno da IA agrupando-os em oito perfis principais: desde os "Pioneiros", que abraçam plenamente a tecnologia, até os "Desafiadores", que se opõem fortemente ao seu uso. A maioria dos pesquisadores (76%) usa a IAG de alguma forma, principalmente para tradução automática (49%) e agentes conversacionais (43%). A aplicação da IAG no ambiente acadêmico foi mais frequente nas etapas de descoberta e edição de pesquisas existentes relatando ganhos em eficiência, especialmente em tarefas de busca e revisão de literatura.

Apesar dos benefícios observados, houve uma desconfiança em relação às empresas de IAG, com apenas 8% dos participantes confiando que essas empresas não usariam seus dados sem permissão. Além disso, 59% dos respondentes acreditam que a IA pode enfraquecer a propriedade intelectual, e 25% acham que reduz a necessidade de pensamento crítico, refletindo preocupações profundas sobre o impacto da tecnologia no rigor acadêmico. As percepções também variam conforme o estágio de carreira, com pesquisadores em início de carreira expressando maior hesitação, enquanto pesquisadores mais experientes tendem a ver a IA como parte essencial do futuro da pesquisa.

Logo, um estudo de planejamento de cenários teve como objetivo visualizar um futuro influenciado pela inteligência artificial (IA) e explorar as incertezas associadas à IA no ecossistema de pesquisa e conhecimento. O processo consultivo, que envolveu mais de 300 pessoas e suas posições em relação ao foco estratégico e às incertezas, destaca pelos cenários extremos. De um lado, a "IA Laissez-faire", caracterizado por oportunidades perdidas, decisões inadequadas e ineficácia, pautado pela incapacidade política e social de se lidar com problemas gerados por tecnologias anteriores à intensificação da IA, como as mídias sociais, e de abordar questões relacionadas a vieses, privacidade, integridade de dados e segurança. Neste cenário, o entusiasmo e a expectativa de que a IA resolveria os problemas mais complexos do mundo resultam em uma adoção precipitada e excessiva da IA. No outro polo, no mundo da "IA Democratizada e Socialmente Integrada" os avanços em interfaces humano-computador, baseados em realidade aumentada aprimorada e outras tecnologias de IA resultariam numa integração de capacidades humanas e computacionais, numa vertente de conhecimento aberto. O design intencional e cuidadoso traria uma convivência fluida, responsável e segura, visando objetivos sociais amplos e bens públicos, impulsionada pela academia e por outros atores sociais inovadores e participantes (ALR/CNI, 2024).

Todavia, até o momento da escrita deste guia, as instituições brasileiras diretamente relacionadas à pesquisa acadêmica[15], a exemplo de MEC, CAPES e CNPq e de agências estaduais de fomento à pesquisa, ainda não produziram quaisquer regramentos ou diretrizes para tais usos no Brasil[16]. Portanto, parece haver uma notável lacuna na

15. A título de exemplo, o Tribunal de Contas da União lançou um guia de uso de inteligência artificial generativa por seus servidores, apresentando uma série de diretrizes e boas práticas. Disponível em: https://portal.tcu.gov.br/data/files/42/F7/91/4B/B59019105E366F09E18818A8/Guia%20de%20uso%20de%20IA%20generativa%20no%20TCU.pdf. Acesso 14 out. 2024.

16. A biblioteca eletrônica SciELO produziu um regramento para seus periódicos. https://wp.scielo.org/wp-content/uploads/Guia-de-uso-de-ferramentas-e-recursos-de-IA-20230914.pdf#:~:text=O%20guia%20estabelece%20normas%20e%20pr%C3%A1ticas%20que%20ser%C3%A3o%20aplicadas%20a. Acesso 2 out. 2024.

pesquisa científica brasileira[17].

Diante deste cenário, este guia foi produzido com o objetivo de orientar pesquisadores e pesquisadoras sobre princípios gerais para o uso ético e responsável de ferramentas de Inteligência Artificial, notadamente, as generativas. Ele foi construído com base na pesquisa dos autores e em uma busca ativa por materiais similares, seja na forma de artigos acadêmicos brasileiros e internacionais que discutem sobre efeitos das IAG na pesquisa acadêmica, mapeiam ou mesmo sugerem melhores práticas, seja no formato de diretrizes presentes em associação científicas, universidades, editoras acadêmicas, periódicos de alto impacto e afins. Em vários momentos, entretanto, as sugestões se baseiam nas próprias experiências dos autores em bancas, congressos, mesas-redondas e interação com os pares, buscando sempre que possível alicerçar as sugestões em base de literatura acadêmica.

Este guia, dessa forma, não busca substituir regramentos de cada universidade, instituição de pesquisa ou mesmo de instituições superiores de fomento à pesquisa, mas auxiliar diretamente pesquisadores que estão se sentindo inseguros e sem devido respaldo, neste momento. Logo, não se trata de regras, mas de diretrizes, ou simplesmente, sugestões de boas práticas para o uso de tais soluções, assim como questões para reflexões. Portanto, regramentos, já existentes ou futuros, publicados por instâncias superiores e por eventuais locais de publicação do material acadêmico (como congressos, editoras e periódicos), certamente deverão ser seguidos no caso de qualquer divergência quanto aos princípios aqui apresentados.

Além disso, cada pesquisador ou pesquisadora deverá ter um olhar crítico para o seu projeto, sua área de pesquisa e as necessidades específicas de sua pesquisa, a fim de garantir que o máximo de esforço foi realizado para se alcançar um uso ético e responsável da IA. Este manual foi produzido por professores da área das Ciências Humanas e Sociais Aplicadas, podendo não refletir adequadamente questões específicas das áreas das Ciências Exatas e da Terra, Ciências Biológicas, Engenharias, Ciências da Saúde, da Vida e das Exatas, entre outras.

Ainda mais, este guia foi especialmente pensado para aplicativos de Inteligência Artificial Generativa e outras soluções que não exigem conhecimentos de linguagem de programação (*no code* ou *low code*), como R e Python. Nosso foco são ferramentas generalistas, como o ChatGPT, ou especializadas para a pesquisa que possam auxiliar em diversas funções do fazer científico, como busca e seleção de literatura, leitura e estudo, escrita acadêmica, análise de dados, apresentação de resultados e tradução. Geralmente, trata-se de ferramentas, sistemas e soluções baseados em sites, plataformas, aplicativos, que produzem resultados através de entradas e comandos, intermediados por interfaces de usuário de uso simples. Por conseguinte, boa parte das sugestões não se destinam a pesquisadores e cientistas que produzem seus próprios códigos e modelos de Inteligência Artificial, nos quais os pesquisadores têm pleno controle da solução e de seus limites.

Finalmente, o cenário da IA Generativa vem se alterando em ritmo acelerado, nos dois anos desde o lançamento do ChatGPT. Tanto as *Big Techs,* diretamente envolvidas no desenvolvimento e lançamento de modelos cada vez maiores e mais potentes, como uma cultura de empreendedorismo por meio de *startups* em busca de aproveitar o *boom* de financiamentos, contribuem para um cenário extremamente dinâmico. Portanto, este guia pode se tornar "datado" em determinados aspectos, notadamente em termos de funções e ferramentas (que podem inclusive deixar de existir!), cabendo ao leitor fazer essa avaliação em cada caso. Dentro do possível, este guia será atualizado.

17. Uma excelente exceção é o dossiê publicado na edição 28 de 2023 da revista TECCOGS – Revista Digital de Tecnologias Cognitivas da PUC-SP sob supervisão da professora Lucia Santaella no formato de um "Manual Ético para o uso da Inteligência Artificial Generativa". Os vários artigos presentes neste manual são bastante úteis para a compreensão de inúmeros aspectos que tais tecnologias impactam no atual momento. Em especial, o artigo de Diego Franco, Luís Eduardo Viegas, e Anderson Röhe (2023) apresenta um excelente guia para o uso ético da inteligência artificial, que iremos reproduzir ao final deste manual como referência extra de boas diretrizes a serem seguidas. O dossiê pode ser acessado em: PUC-SP. Manual Ético para o uso da Inteligência Artificial Generativa. **TECCOGS - Revista Digital de Tecnologias Cognitivas**, São Paulo, n. 28, 2023. Disponível em: https://revistas.pucsp.br/index.php/teccogs/issue/view/2973/495. Acesso em: 30 set. 2024.

Princípios gerais

Inicialmente, é importante apresentar e destacar os princípios gerais para o uso de ferramentas de Inteligência Artificial Generativa (IAG), orientados a questões mais amplas e a e princípios normativos para orientar o uso responsável e ético. Tais práticas e reflexões devem sempre atravessar todos os usos e regramentos de IAG na pesquisa acadêmica.

I. Compreensão das ferramentas de IAG

Antes de adotar a Inteligência Artificial Generativa em seus processos de trabalho, **pesquisadores devem compreender adequadamente os termos de serviço, as políticas de privacidade e as implicações de segurança associadas a essas tecnologias.** Esta compreensão é um passo importante para preservar a integridade acadêmica e garantir o uso ético e responsável da IA (Franco *et al*., 2023; Unesco, 2024; União Europeia, 2024).

Antes de tudo, devemos nos lembrar que os Grandes Modelos de Linguagem e outras soluções de IAG não são desenhados especificamente para a pesquisa científica, sendo propriedade de grandes corporações, usualmente denominadas como *Big Techs* (Morozov, 2018). Como elas funcionam à base de gigantescos financiamentos, elas precisam garantir o lucro e se manterem no topo das disputas, apresentando os melhores modelos e produtos, no que tem sido considerado a corrida pela IA (*AI race*). Isso frequentemente implica no lançamento de produtos inacabados e que não atendem a todos aspectos de segurança, o que torna ainda mais complexo a sua utilização acadêmica (Liao e Vaughan, 2023, p. 5).

Considerando tal necessidade de viabilidade financeira, a melhor maneira de manter essas soluções gratuitas ou a valores módicos consiste em um padrão semelhante ao que vivenciamos nas redes sociais digitais. Como pagamento, os usuários entregam seus dados para as empresas, que fazem uso deles para melhorar seus produtos (Morozov, 2018)[18].

Portanto, os dados inseridos em tais tecnologias podem ser utilizados e potencialmente retidos pela plataforma, tornando-se parte de conjuntos de dados futuros, isto é, que serão usados para o treinamento dos modelos. Neste sentido, uma subsidiária da gigante editora de publicação acadêmica Taylor & Francis, a Informa teria assinado um contrato para fornecimento de material de treinamento para a Microsoft, sem o consentimento dos autores de artigos, capítulos e livros[19]. Logo, cabe aos pesquisadores avaliarem criticamente as políticas e práticas da empresa proprietária da plataforma. *Para maior precisão, esses termos devem ser acessados fora da plataforma de IA, em vez de consultar a própria IA, que pode simplesmente inventar uma resposta* (Liao, Vaughan, 2023; Gao *et al*., 2024).

Os pesquisadores devem estar cientes das fontes de dados usadas no treinamento da IA, considerando possíveis vieses ou limitações decorrentes desses materiais. A falta de explicação sobre como os resultados são gerados tende a impor aos usuários a lógica definida pelos parâmetros projetados nos sistemas de IA generativa (Silva *et al*., 2024). Os modelos atuais, como o ChatGPT, são treinados com dados on-line que refletem predominantemente os valores e as normas do Norte Global, tornando-os potencialmente inadequados para comunidades carentes de dados em várias partes do Sul Global ou em comunidades mais desfavorecidas do Norte Global (Sampaio *et al*., 2024 a; Unesco, 2024).

No atual momento, a possibilidade de criar e controlar a IA Generativa está fora do alcance da maioria das empresas e países, especialmente aqueles do Sul Global. Isso pode levar a uma amplificação da desigualdade de acesso a recursos, particularmente em países em desenvolvimento, onde já se observa notáveis desigualdades digitais (Dwivedi *et al.*, 2023; Rikap, 2021; Unesco, 2024). Além disso, é importante notar que tais Grandes Modelos de Linguagem operam em escala global, e suas respostas podem não estar alinhadas com leis e regulamentações locais (Dwivedi *et al*., 2023).

É fundamental que os pesquisadores conheçam os direitos de propriedade dos dados e verifiquem se as ferramentas de IA generativa que estão utilizando

18. O que tende a alimentar uma lógica de colonialismo de dados. Ver a seção final deste documento para uma discussão mais aprofundada sobre a questão.

19. POTTER, Wellett. An academic publisher has struck an AI data deal with Microsoft – without their authors' knowledge. **The Conversation**, 23 jul. 2024. Disponível em: https://theconversation.com/an-academic-publisher-has-struck-an-ai-data-deal-with-microsoft-without-their-authors-knowledge-235203. Acesso em: 2 out. 2024.

violam alguma regulamentação brasileira vigente, notadamente a Lei Geral de Proteção de Dados[20]. Também devem estar cientes de que imagens, áudios, vídeos ou códigos criados com IA generativa podem violar direitos de propriedade intelectual de terceiros, e que o conteúdo que eles criam e compartilham na Internet pode ser explorado por outras IAs generativas (Gao *et al.*, 2024; Unesco, 2024).

Os pesquisadores devem, ainda, tomar cuidado para não fornecer dados pessoais de terceiros a sistemas de IA Generativa, a menos que o titular dos dados tenha dado seu consentimento. É preciso que os pesquisadores tenham um objetivo claro para o qual os dados pessoais serão usados, garantindo a conformidade com as regras de proteção de dados do local da pesquisa (Bail, 2024; Ray, 2023; Hoseil *et al*., 2023; União Europeia, 2024; Unesco, 2024).

Para mitigar riscos, recomenda-se evitar compartilhar dados confidenciais, proprietários ou com implicações de propriedade intelectual ao usar ferramentas de IA generativa. Alternativamente, pode-se optar por usar apenas modelos abertos[21] sob controle do pesquisador ou aplicações em que haja certeza de que as informações não são retidas ou utilizadas pela empresa para treinamento de seus modelos (Bail, 2024; Limongi, 2024), conforme será abordado posteriormente.

É importante também compreender que diferentes soluções de IA podem ser baseadas nos Grandes Modelos de Linguagem, mas apresentar regras diferentes dos provedores originais. Logo, diferentes aplicações para busca, geração de texto, imagens e apresentações, análises de dados e afins rodam com base nos modelos, mas podem ter políticas de privacidade, segurança e uso de dados diferente da plataforma original (Liao e Vaughan, 2023). Em um exemplo simples, durante a escrita deste texto, era possível desativar o treinamento do modelo no ChatGPT, mas não é possível no uso de modelos no *Poe*. Por sua vez, o uso do modelo do ChatGPT para análises qualitativas no Atlas.ti é protegido (Velez, Rister, 2024), portanto, tais verificações das normas devem ser realizadas em cada aplicativo.

II. Autoria humana

Um dos poucos consensos já estabelecidos no atual momento é de que ferramentas de IAG, como ChatGPT, Copilot, Claude, Gemini, Maritalk e outros Grandes Modelos de Linguagem, não podem ser listadas como autores, pois **a autoria exige uma pessoa legal e responsável (*accountable*) pelo conteúdo** (Thorp, 2023; Stokel-Walker, 2023). IAs são incapazes de assumir responsabilidade moral ou legal pela originalidade, precisão e integridade do trabalho. Somente humanos podem garantir que o conteúdo reflita as ideias dos autores e esteja livre de plágio, fabricação ou falsificação, incluindo textos e imagens gerados pela IA.

A responsabilidade e a prestação de contas (*accountability*) são essenciais para garantir o uso ético da IA na pesquisa. Isso inclui o estabelecimento de diretrizes claras para a conduta ética em pesquisas que utilizam IA, bem como mecanismos para monitorar e fazer cumprir essas diretrizes (Franco *et al*., 2023; Limongi, 2024). Os autores devem assumir total responsabilidade pela integridade do conteúdo gerado pela IA, incluindo a revisão e a edição cuidadosa para evitar informações e citações incorretas, incompletas, inventadas ou tendenciosas. Portanto, **a última aprovação da versão final do produto acadêmico a ser publicado é sempre uma tarefa humana** (Cambridge, 2023; COPE, 2023; Elsevier, 2023b; ICMJE, 2023; Oxford, 2023; Taylor & Francis, 2023; Wiley, 2023; Zielinski *et al*., 2023; Perkins e Roe, 2024).

Em especial, as recomendações da editora Sage (2023) explicitam questões como a perpetuação de vieses e estereótipos existentes nos dados de treinamento, sendo dever dos autores avaliar seu efeito nos resultados. Da mesma maneira, deve-se checar se os conteúdos gerados ou revisados por essas ferramentas estão livres de plágio, afinal elas podem reproduzir texto de outras fontes existentes em seu banco de dados, como veremos em mais detalhes abaixo. Recomenda-se, então, um monitoramento contínuo e participação humana em todas as fases da pesquisa (Franco *et al*., 2023; Unesco, 2024; União Europeia, 2024).

III. Transparência

Quando tratamos da temática, é importante uma diferenciação: podemos falar de transparência dos modelos e aplicações de IAG e podemos falar da transparência do uso pelos pesquisadores. Ambas são importantes, entretanto a primeira é particularmente desafiadora e não está nas mãos dos pesquisadores.

20. Recomendamos a leitura do documento produzido pela Autoridade Nacional de Proteção de Dados sobre "Tratamento de dados pessoais para fins acadêmicos e para a realização de estudos e pesquisas" (ANPD, 2023).

21. O Hugging Face é uma interessante iniciativa da promoção e aprimoramento de vários modelos abertos de inteligência artificial. Disponível em: https://huggingface.co/. Acesso em: 4 out. 2024.

Via de regra, estamos discutindo tecnologias proprietárias desenvolvidas por grandes corporações, disputando diretamente o novo mercado de IAG; como consequência, como os modelos funcionam é uma questão protegida por leis de segredo industrial. Com poucas informações sobre os gigantescos bancos de dados utilizados, incluindo as suas fontes distintas, além de ainda menos detalhes sobre os aspectos técnicos dessas tecnologias, devemos considerar os Grandes Modelos de Linguagem como caixas-pretas, cabendo à comunidade acadêmica avaliar *outputs* dos modelos e pressionar as empresas por maior transparência em seu funcionamento (Liao, Vaughan, 2023, p. 4) .

Tendo esta consideração em vista, o segundo consenso presente em diferentes diretrizes sobre o uso de IAs generativas é a necessidade de transparência sobre a utilização de tais soluções na pesquisa acadêmica. *Pesquisadores que fizerem uso de Inteligência Artificial Generativa devem descrever na coverletter e no manuscrito como utilizaram a ferramenta para garantir transparência, replicabilidade e confiabilidade da pesquisa* (Cambridge, 2023; COPE, 2023; ICMJE, 2023; Oxford, 2023; Zielinski *et al*., 2023; Perkins, Roe, 2024; Soares *et al*., 2023).

Para atender a esse requisito de transparência, os pesquisadores devem detalhar quais ferramentas de IAG foram utilizadas substancialmente em seus processos de pesquisa. Isso inclui fornecer informações sobre a ferramenta utilizada, como nome, modelo, versão, e data de uso, além de explicar como ela foi empregada e como afetou o processo de pesquisa (União Europeia, 2024). Tal transparência deve sempre buscar permitir ao máximo a replicabilidade da pesquisa (Liao, Vaughan, 2023; Resnik, Hosseini, 2024; Silva, Bonacelli, Pacheco, 2024).

Como norma geral, isso deve incluir a disponibilização dos *prompts* e dos resultados advindos (*outputs*) do modelo, como os textos, os *scripts* de programação, visualizações e mesmo resultados de análise[22]. Caso seja possível, incluir também o *link* com o *log* (a conversa completa no chat) da interação com a IA. No caso do uso de API[23], tais resultados podem ser tornados disponíveis como materiais complementares. Para certas atividades de pesquisa, é aconselhável fornecer mais detalhes do que um simples reconhecimento da ferramenta utilizada, como as escolhas dos diferentes *prompts* feitos para obter os resultados (Sampaio *et al*., 2024a; Schulhoff *et al*., 2024 Susarla *et al*., 2023).

Os pesquisadores devem levar em conta a natureza estocástica (aleatória) das ferramentas de IA generativa, que tendem a produzir diferentes resultados a partir do mesmo insumo (*input*) ou pedido (*prompt*). Portanto, devem visar a reprodutibilidade e robustez em seus resultados e conclusões, divulgando e discutindo as limitações das ferramentas de IAG utilizadas, incluindo possíveis vieses no conteúdo gerado, bem como possíveis medidas de mitigação (União Europeia, 2023). Como tais modelos são atualizados com muita frequência, é importante incluir também as datas das consultas. Uma abordagem transparente não apenas promove a confiança na comunidade científica, mas também permite uma avaliação mais precisa do papel da IA no processo de pesquisa e seus possíveis impactos nos resultados obtidos.

Todas essas sugestões indicam para a explicitação do uso de IAG, por parte dos autores. Algumas editoras e revistas sugerem a inclusão de uma declaração ao final do manuscrito (Elsevier, 2023b. Resnick), imediatamente antes das referências, com o título "*Declaração de IA e tecnologias assistidas por IA no processo de escrita*". Nela, os autores devem especificar a ferramenta utilizada, o motivo e a forma de aplicação, o motivo e a forma de aplicação, evidenciando cada etapa da pesquisa que teve o uso de IA. Abaixo, um exemplo.

Formato Sugerido para a Declaração:

"Durante a preparação deste trabalho, o(s) autor(es) utilizou(aram) [nome da ferramenta/modelo ou serviço] versão [número e/ou data] para [justificar o motivo]. Após o uso desta ferramenta/modelo/serviço, o(s) autor(es) revisou(aram) e editou(aram) o conteúdo em conformidade com o método científico e **assume(m) total responsabilidade pelo conteúdo da publicação**".

A adoção de modelos de IA de código aberto é uma

22. Tal recomendação, reconhecemos, pode soar utópica em determinados contextos. Embora a transparência esteja de acordo com os imperativos mertonianos do *ethos* científico, na prática observamos o "segredismo", principalmente quando a metodologia é um diferencial competitivo.

23. Uma API (*Application Programming Interface*, em inglês, ou Interface de Programação de Aplicações, em português) é uma conexão entre sistemas que permite trocar informações e serviços. Para nossos interesses, você pode acionar modelos de linguagem por APIs e aí não terá a interação tradicional pela janela do chatbot.

maneira eficaz de promover a transparência. Tais modelos facilitam a revisão e colaboração entre pesquisadores e permitem auditorias públicas e independentes, facilitando que a IA seja usada de maneira justa e ética (Bail, 2024; Limongi, 2024), como voltaremos a debater ao longo deste material.

IV. Integridade da pesquisa acadêmica

Desde sua difusão pública, a IAG tem suscitado desconfiança e resistência por parte do campo educacional e acadêmico. O motivo desta atitude negativa se relaciona com uma percepção de uso indevido, contrário às normas formais e informais que situam claramente o papel da autoria e o valor da originalidade na academia. Assim, o uso da IAG na pesquisa científica apresenta desafios significativos para a integridade acadêmica e alerta para a fraude e para comportamentos antiéticos[24].

Conforme o documento "Diretrizes para a ética na pesquisa e a integridade científica", elaborado pelo Grupo de Trabalho de Ética em Pesquisa do Fórum de Ciências Humanas, Sociais, Sociais Aplicadas, Linguística, Letras e Artes (FCHSSALLA, 2024, p. 11), a integridade acadêmica e científica "é um dos pilares da prática científica e consiste no compromisso com a construção coletiva da ciência, de forma transparente, responsável, rigorosa e honesta". A comunidade deve ativamente "coibir e combater a falsificação, a fabricação de dados e o plágio, consideradas como as três violações mais graves à integridade científica" (*Ibidem*, p. 15).

Diferente de modelos tradicionais de aprendizado de máquina, nos quais temos estruturas mais típicas de entradas e saídas (*inputs* e *outputs*), os Modelos de Linguagem são consideravelmente mais flexíveis, podendo ser usados para diversas tarefas de pesquisa, notadamente aquelas ligadas a texto: responder a perguntas, gerar diálogos, completar frases, resumo, parafraseio, elaboração, escrita e mesmo classificação e análise (Liao, Vaughan, 2023; Silva *et al.*, 2024).

Atualmente, não há respostas estabelecidas, nem mesmo na comunidade de pesquisa, para perguntas, por exemplo, como e por que esses modelos funcionam tão bem quanto funcionam, por que eles podem ou não realizar determinadas tarefas e como as características dos dados de treinamento afetam as capacidades do modelo (Liao, Vaughan, 2023, p. 4, tradução nossa).

Dito de outra forma, nem mesmo os especialistas e desenvolvedores dos sistemas de IAG conseguem explicar completamente como os resultados desses modelos são gerados. Como mencionado anteriormente, são máquinas de correlações estatísticas que não "raciocinam" como seres humanos (Kaufman, 2002). Assim, esses sistemas de aprendizado de máquina são incapazes de elucidar os mecanismos causais que produzem seus resultados, deixando uma parte obscura até mesmo para seus criadores. "Pela primeira vez na história, criamos máquinas que operam de uma forma que seus criadores não entendem [...], um instrumento capaz de aplicar uma nova lógica de conhecimento" (Silva *et al*., 2024, p. 755, tradução nossa).

Isso, claro, não significa que não existam uma série complexa de parâmetros técnicos para os diferentes modelos. Justamente, a falta de transparência nos algoritmos e critérios utilizados pelas IAG pode levar a uma compreensão limitada sobre como as decisões são tomadas. Isso é particularmente problemático quando se trata de recomendações de leituras, métodos estatísticos ou representações visuais na ciência. Um dos principais riscos é a possibilidade de as IAs produzirem respostas que, embora pareçam plausíveis, podem ser descontextualizadas, factualmente incorretas ou distorcidas pelos vieses do modelo (Liao, Vaughan, 2023; Rahman *et al*., 2023; Cong-Lem et al., 2024; Khan *et al*., 2024).

Dito de outra forma, as IAGs podem replicar, e possivelmente, perpetuar vieses presentes em seus dados de treinamento, o que pode influenciar os resultados da pesquisa de maneiras sutis e difíceis de detectar. A falta de controle do pesquisador sobre muitos aspectos dessas tecnologias torna desafiador mensurar a validade e confiabilidade dos resultados gerados (Grossi *et al*., 204; Ramos, 2023; Ray, 2023; Susarla *et al*., 2023).

A inconsistência nas respostas da IA generativa é uma preocupação adicional. A mesma análise pode produzir resultados diferentes em momentos distintos, o que complica ainda mais a confiabilidade da pesquisa baseada nessas ferramentas. Em suma, embora a IAG ofereça potenciais benefícios para a pesquisa científica, seu uso requer uma supervisão humana rigorosa, uma compreensão clara de suas limitações e uma abordagem crítica constante para garantir a integridade e a qualidade da pesquisa acadêmica (Barreto, Ávila, 2023; Limongi, 2024; Sampaio *et al.*, 2024a).

24. Para uma discussão "pré-GPT" sobre estes conceitos, assim como práticas associadas, a exemplo do autoplágio e da "ciência salame", entendidas num contexto de produtivismo acadêmico, conferir Sabbatini (2013).

A natureza dinâmica dos Grandes Modelos de Linguagem também apresenta desafios para a replicabilidade da pesquisa, pois diferentemente de *softwares* acadêmicos que mantêm versões estáveis e rastreáveis, os LLMs são frequentemente atualizados sem preservar o acesso às versões anteriores (falta de rastreabilidade ou *traceability*) e frequentemente tais mudanças não estão facilmente disponíveis para pesquisadores. Isso pode tornar certos aspectos da pesquisa complemente não replicáveis (Dwivedi *et al*., 2023; Liao, Vaughan, 2023; Sampaio *et al.*, 2024a).

Inicialmente, pesquisadores devem identificar áreas onde a IA pode agregar valor sem comprometer a integridade do trabalho acadêmico. Isso tende a naturalmente incluir o uso de IA para tarefas repetitivas ou de organização, buscando liberar tempo para análise crítica e pensamento original (Kaufman, 2022; Pereira *et al*., 2023; Ray, 2023; Santaella, 2023; Limongi, 2024)[25].

Em caso de tarefas mais substantivas, pesquisadores devem manter uma atitude crítica em relação ao conteúdo gerado por IAs, adotando estratégias de triangulação e validação cruzada (muitas vezes essa ocorre na comparação com a codificação humana). As informações da IA devem ser verificadas com múltiplas fontes confiáveis, servindo como ponto de partida para pesquisas mais aprofundadas, não como fonte definitiva. Deve-se realizar avaliações regulares da eficácia e do impacto das ferramentas de IA, ajustando ou abandonando aquelas que não atendam aos padrões éticos ou de qualidade (Franco *et al*., 2023; Liao, Vaughan, 2023; Bail, 2024; Soheil *et al*., 2024).

Outra prática importante é manter uma documentação detalhada sobre como as ferramentas de IA foram usadas, incluindo *prompts* específicos e resultados obtidos, registrando inclusive aqueles não utilizados no trabalho final. Tais registros devem ser precisos no caso de inúmeros testes e utilizações, aumentando a transparência do processo de pesquisa, facilitando a replicação e validação dos resultados (Liao, Vaughan, 2023). Essas abordagens críticas e passíveis de adaptação garantem que o uso de IAG na pesquisa permaneça alinhado com princípios éticos e científicos. A verificação rigorosa de erros ou imprecisões no conteúdo gerado é essencial para evitar a perpetuação e amplificação de falhas em pesquisas futuras (Franco, Viegas, Röhe, 2023; Ray, 2023; Santaella, 2023).

Conforme já relatado, uma saída para garantir a integridade e a ética na pesquisa científica que utiliza IA é a adoção de modelos de código aberto como padrão. Esta abordagem não só promove a transparência e a colaboração entre pesquisadores, mas também permite auditorias independentes, fundamentais para o uso justo e ético da tecnologia. Propomos que as instituições de pesquisa e os desenvolvedores de IA priorizem a criação de sistemas hospedadas localmente ou baseadas em nuvem que elas próprias governam. Isso permite que seus colaboradores alimentem seus dados científicos em uma ferramenta que garante a proteção e confidencialidade dos dados. As organizações também devem garantir o nível apropriado de segurança cibernética desses sistemas, especialmente aqueles conectados à Internet. Tais sistemas devem ser explicáveis e auditáveis, além de possibilitarem inúmeros testes para aplicações (Franco *et al.*, 2023; Liao, Vaughan, 2023; Ray, 2023; Limongi, 2024; Unesco, 2024; União Europeia, 2024).

Essas práticas ajudam a garantir que as pesquisas e os resultados baseados em IAG sejam melhor avaliados como transparentes, confiáveis, replicáveis e válidos, observando-se claro os limites de tais ferramentas.

V. Plágio, originalidade e direitos autorais

Os Grandes Modelos de Linguagem treinados em dados disponíveis na Internet podem gerar resultados que infrinjam direitos autorais e, inclusive, caracterizar plágio. Portanto, cabe aos autores buscar garantir a originalidade do trabalho e citar todas as fontes de forma adequada. Além disso, é desejável que os autores tenham os direitos autorais de todo o conteúdo utilizado, incluindo aquele gerado por IA. No momento da escrita deste guia, há um consenso que os resultados (*outputs*) de inteligência artificial generativa não são resguardados por direitos autorais (Samuelson, 2023)[26] .

Ao utilizar conteúdo gerado por IA, “o resultado

25. Ver discussão sobre hypes da IA na seção final deste documento.

26. Há um movimento de empresas de IAG para negociar direitos com empresas produtoras de conteúdo. Confira mais sobre o assunto em: MOHERDAUI, Luciana. Sob pressão por violar direitos autorais, OpenAI faz acordos em série. **Poder 360**, 22 ago. 2024. Disponível em: https://www.poder360.com.br/opiniao/sob-pressao-por-violar-direitos-autorais-openai-faz-acordos-em-serie/. Acesso em: 14 out. 2024.

precisará ser cuidadosamente verificado, já que o *software* parece gerar conteúdo impreciso, baseado em fontes de ideias relatadas de forma imprecisa" (Dwivedi *et al.*, 2023, p. 5), ou ainda, que os resultados espelhem ideias e resultados já publicados por outros autores (Unesco, 2024; União Europeia, 2024, p. 7).

Conforme apresentado em tópicos anteriores, os pesquisadores devem atentar-se às regras das plataformas que fazem uso. Muitas soluções de IA generativa utilizam as interações dos usuários para treinar seus modelos. O site do ChatGPT, por exemplo, deixa claro que "melhoramos continuamente os nossos modelos através de avanços na pesquisa, bem como da exposição a problemas e dados do mundo real". Além disso, é afirmado que o conteúdo compartilhado pelos usuários pode ser utilizado para esse fim: "Quando você compartilha seu conteúdo conosco, isso ajuda nossos modelos a se tornarem mais precisos e melhores na solução de seus problemas específicos, além de ajudar a aprimorar os recursos gerais e a segurança deles"[27]. Portanto, sob pena de tornarem públicos dados de pesquisa, pesquisadores devem evitar carregar trabalhos não publicados ou sensíveis em sistemas de IA online, a menos que haja garantias de que os dados não serão reutilizados. Da mesma forma, frequentemente os resultados advindos da IAG podem estar mimetizando ideias já elaboradas anteriormente (Unesco, 2024, União Europeia, 2024).

Evidentemente, parte das discussões advindas da rápida adoção das soluções de IAG está na capacidade ou não de outras ferramentas detectarem o conteúdo produzido por máquinas, a exemplo do que já havíamos inserido em nosso cotidiano de instrumentos capazes de detectar indícios de plágio. As empresas de detecção de plágio estão enfrentando dificuldades para abordar a questão específica das IAG. Apesar de muitas empresas reivindicarem terem sistemas capazes de detectar a geração de textos pelas IAs, são frequentes as reclamações de instituições de ensino e pesquisa apontando o contrário[28]. Portanto, no momento da redação deste manual, acreditamos que o uso desses instrumentos deve vir acompanhado de especial cautela, sob pena de termos pesquisadores indevidamente penalizados por plágio, enquanto outros permanecem não detectados (Dwivedi *et al*., 2023, p. 27).

Não há caminhos fáceis a serem seguidos. Conforme Cotton *et al.* (2023), precisamos de estratégias combinadas de técnicas automatizadas e manuais (humanas) de avaliação, na qual professores e pesquisadores sejam resguardados para fazer avaliações sobre o uso de IA de forma inadequada. Todavia, Liao e Vaughan (2023, p. 5, tradução nossa) nos lembram da importância do modelo mental das pessoas, ou ainda, "a percepção dos usuários sobre o que o modelo ou sistema pode ou não fazer e como ele funciona". Em outras palavras, os pesquisadores e professores podem ter uma visão distorcida sobre tais tecnologias se não as compreenderem adequadamente. Um exemplo disso é a prática de usar o próprio modelo para tentar detectar plágio, o que não tem absolutamente qualquer confiabilidade[29].

Outra preocupação relacionada é a proliferação de artigos de baixa qualidade ou plagiados, bem como a erosão da confiança na comunidade acadêmica. Também pode resultar na desvalorização das habilidades e conhecimentos necessários para produzir publicações acadêmicas de qualidade (Cotton *et al.*, 2023; Curtis, 2023; Velez, Rister, 2024).

Logo, também devem existir estratégias para educar estudantes e jovens pesquisadores sobre o plágio. Ao avaliar as políticas de uso de IA por universidades americanas, Velez e Rister (2024) denotam que essa discussão pode acontecer de forma colaborativa entre professores e estudantes, podem levar a críticas e reflexões produtivas sobre voz, estilo e estrutura, assim como a IA pode ajudar os estudantes a entenderem seus papeis como autores, e editores. Veja mais sobre isso no tópico de detecção.

É importante ressaltar que um dos principais critérios para publicação em um periódico acadêmico é a novidade. A IA Generativa ainda não pode igualar a originalidade humana. Dessa

27. No caso da OpenAI e de outras empresas como a Meta, o usuário pode optar por não treinar o modelo por meio de configurações ou preenchendo um formulário. O ChatGPT, especificamente, oferece uma opção de chat privado, que não treina o modelo de IA. Sempre busque verificar se o modelo usa seus dados para seu treinamento e se existem possibilidades privadas ou mesmo de não permitir que isso seja realizado antes de subir dados sensíveis, inéditos ou protegidos.

28. Confira: COLEY, Michael. Guidance on AI Detection and Why We're Disabling Turnitin's AI Detector. **Vanderbilt University**, 16 ago. 2023. Disponível em: https://www.vanderbilt.edu/brightspace/2023/08/16/guidance-on-ai-detection-and-why-were-disabling-turnitins-ai-detector/. Acesso em: 23 out. 2024.

29. LOBATO, Gisele. Uso do ChatGPT gera conflitos na sala de aula e acusações de plágio sem provas. **Aos Fatos**, 10 jul. 2023. Disponível em: https://www.aosfatos.org/noticias/chatgpt-alunos-professores-plagio/. Acesso em: 23 out. 2024.

forma, o principal uso para a IA generativa a curto prazo ainda deve ser como assistente de pesquisa (Dwivedi *et al.*, 2023, p. 41).

VI. Preservação da agência humana

A ideia de humanos e máquinas trabalhando juntos é tão antiga quanto o próprio campo da IA, na qual o objetivo de tais ferramentas seria aumentar a inteligência humana para resolver problemas complexos e tomar melhores decisões (Dwivedi *et al*., 2023)[30]. Assim, o material gerado por IAG pode parecer preciso e convincente, mas muitas vezes contém erros ou ideias e resultados tendenciosos. Isso representa um alto risco para estudantes e pesquisadores jovens que não possuem conhecimentos prévios sólidos sobre o tópico em questão.

É necessário adotar uma perspectiva crítica sobre tudo o que eles "produzem". A criação de conhecimento envolve um processo rigoroso que pode incluir teorização, teste de hipóteses, coleta de dados, análise, interpretação e avaliação cuidadosa, algo que a IA Generativa sozinha ainda não pode replicar (Susarla *et al*., 2023).

Para salvaguardar a agência humana, as instituições de ensino e pesquisa devem priorizar o desenvolvimento de habilidades para a pesquisa, usando a IAG como uma ferramenta complementar, não como substituta do aprendizado. Para um jovem pesquisador, por exemplo, uma parte importante do processo de aprendizagem é adquirir habilidades de análise, síntese e discussão dos resultados de pesquisa. Usar a IAG para escrever esse tipo de conteúdo pode privar o estudante da oportunidade de desenvolver essas habilidades (Alves, 2023; Gonçalves, 2023; Unesco, 2024; União Europeia, 2024; Habib *et al*., 2024).

A dependência excessiva da IAG pode levar a experiências de aprendizagem incompletas, o que é especialmente relevante se o caráter formativo do jovem pesquisador for levado em conta. Assim, o uso responsável da IA e a conscientização de suas limitações deve estar em consonância com processo de crescimento e de alcance da autonomia humana. A independência intelectual, o pensamento crítico e mesmo a criatividade do pesquisador em formação passam pela consideração da tecnologia em relação a seu próprio enfoque de pesquisa, em seu papel de ferramenta de apoio - não de substituição de seu esforço pessoal (Cotton *et al*., 2023; Franco *et al*., 2023; Velez, Rister, 2024).

Portanto, é preciso evitar antropomorfizar as IAs generativas. Elas não são seres humanos ou mesmo seres sencientes[31]. Não devemos tratá-las como substitutos de parceiros acadêmicos, orientadores e afins. Elas não podem servir para minar as relações humanas, que já tendem a ser um problema no mundo acadêmico. No atual momento, elas são inclusive propensas a sempre concordar com o usuário e isso pode levar a caminhos menos produtivos em termos científicos e mesmo de saúde mental. A pesquisa acadêmica é naturalmente difícil e complicada em vários momentos e se trata de um campo para ampla discussão e divergências (Cruz, 2020). As IAs não podem se tornar suportes para evitar frustrações na vida acadêmica (Alves, 2023; Gonçalves, 2023). O segredo aqui é entender o papel delas como ferramentas de pesquisa e não como parceiras ou colegas no sentido estrito.

Também é importante chamar a atenção para o conceito de "*human in the loop*" (humano no circuito, em tradução do inglês), isto é, a assertiva de que os humanos desempenham um papel central, supervisionando e interagindo com os sistemas de IA, mantendo o controle final sobre as decisões e saídas geradas. Diferentemente de uma tecnologia operando de forma autônoma, e especialmente diante da capacidade criativa, a supervisão e intervenção humanas são necessárias para garantir a relevância, segurança e alinhamento ético das saídas produzidas pelo sistema. Uma abordagem "*human in the loop"*, como consequência de sua própria definição, age em sentido contrário à substituição do humano pela máquina.

A frase de Tudor Jones, "Nenhum homem é melhor do que uma máquina, mas nenhuma máquina é melhor do que um homem com uma máquina", reforça o conceito de *human in the loop*, pois destaca a sinergia entre a capacidade humana e a eficiência da máquina. Enquanto a máquina proporciona velocidade e precisão, o papel humano é fundamental para direcionar essa eficiência, aplicando julgamento crítico, intuição e valores éticos às saídas automatizadas. Essa colaboração é especialmente relevante no contexto de sistemas de IA em ambientes complexos, como o acadêmico, onde o valor da máquina está em potencializar a análise e a síntese, mas onde o discernimento humano garante a relevância e integridade das

30. "Amplified Cognition", "Augmented Intelligence", "Collaborative Intelligence", "Extended Intelligence" etc.
31. Para uma aprofundada e interessante discussão sobre o assunto, ver Röhe e Santaella (2023).

conclusões. Assim, a frase reflete a importância de aliar as vantagens operacionais da tecnologia com a supervisão humana para atingir resultados robustos e eticamente alinhados.

Neste sentido, no atual momento, o desenvolvimento de habilidades de engenharia de *prompts* é fundamental, ou ainda, as habilidades para saber fazer boas perguntas e obter boas respostas dos modelos de IAG. A Unesco (2024) recomenda utilizar uma linguagem clara e direta, incluir exemplos e contexto, e refinar as consultas conforme necessário. Esta abordagem não apenas melhora a qualidade das respostas obtidas, mas também desenvolve habilidades críticas de interação com sistemas de IA. No entanto, é importante notar que estudantes mais jovens podem não ter a paciência necessária para interagir com o modelo até obter a melhor resposta. Há diversos sites e repositórios de *prompts* (ver Apêndice 2), além de recomendações para *prompts* científicos (Bsharat *et al*., 2023, Girardi, Pase, 2024; Jesus, Santatém Segundo, 2024; Schulhoff *et al*., 2024).

Em última análise, o objetivo é usar a IAG como uma ferramenta para transformar a pesquisa, ajudando os estudantes a adquirir habilidades de aprendizagem ao longo da vida e a usá-las em suas futuras carreiras para resolver problemas reais em suas tarefas. No entanto, deve-se equilibrar esse uso com o desenvolvimento contínuo das habilidades de pensamento independente e expressão linguística dos futuros pesquisadores, incentivando-os a avaliar e ser críticos em relação a todo tipo de resultado (*output*) gerado pelas IAG.

VII – Uso eticamente orientado

Ainda que este guia tenha como foco o uso ético das soluções de Inteligência Artificial Generativa, desejamos enfatizar que as considerações éticas devem permear todo o processo de pesquisa envolvendo IAG, desde a concepção até a publicação dos resultados. A integração de IA generativa na pesquisa acadêmica demanda uma abordagem cuidadosa e eticamente orientada, que vai além da mera avaliação de qualidade e eficácia dos sistemas. Demanda, portanto, uma reflexão constante sobre as implicações éticas, considerando seu potencial impacto nos participantes da pesquisa e na sociedade em geral.

Para facilitar, o texto elaborado pelo FCHSSALLA (2024, p. 12) apresenta os princípios orientadores para uma pesquisa ética, nomeadamente:

a. respeito à liberdade, à igualdade, à dignidade e à autonomia das pessoas e a todas as formas de vida;
b. respeito à diversidade cultural, social, religiosa, étnico-racial, linguística, geracional, territorial, moral, sexual e de gênero;
c. respeito às características e às necessidades das pessoas com deficiência;
d. responsabilidade na condução e na execução da pesquisa;
e. independência e autonomia na realização da pesquisa;
f. compromisso com a integridade acadêmica e com a honestidade intelectual;
g. diálogo permanente com a comunidade científica e com a sociedade;
h. empenho na divulgação do conhecimento em veículos e formatos acessíveis;
i. transparência em todas as atividades acadêmicas e científicas;
j. responsabilidade no uso de recursos financeiros da pesquisa.

Conforme este documento, dentre outros aspectos, os sujeitos de pesquisa têm o direito de serem abordados de maneira respeitosa; têm a garantia do respeito à sua privacidade e identidade; de manifestar o seu consentimento ou não consentimento; de serem informadas sobre implicações, riscos e danos decorrentes da participação na pesquisa e sobre estratégias para evitá-los. Nenhum ponto desses deve ser ignorado ao fazer uso de IAs, notadamente quando suas identidades e dados sensíveis podem ser expostos. É essencial adotar uma abordagem de IA centrada no ser humano (ver tópico anterior), alinhada com os princípios dos direitos humanos, que proteja a dignidade humana e respeite a diversidade cultural.

Para promover um uso ético e responsável de IA na pesquisa, as instituições acadêmicas devem estabelecer diretrizes claras e mecanismos de supervisão. É necessário monitoramento contínuo de usos e resultados de IAG, o que sugere a formação de comitês diversificados e transparentes, representativos da pluralidade acadêmica[32], incluindo educadores, pesquisadores, programadores e engenheiros de IA e representantes de outras partes interessadas, como grupos que são geralmente mais afetados

32. Um bom exemplo vem da UFMG que já montou um comitê permanente de inteligência artificial. Disponível em: http://www.ufmg.br/ia/. Acesso em: 2 out. 2024.

por tais tecnologias, como mulheres, negros, LGBTQIAPN+, povos indígenas, além de outros grupos marginalizados de forma geral (Silva, 2022; Franco *et al*., 2023; Stahl, Eke, 2024; Unesco, 2024). O objetivo é otimizar o potencial da IA Generativa na educação e na pesquisa, ao mesmo tempo em que se implementam medidas eficazes para mitigar os riscos associados à sua adoção (Franco *et al*., 2023; Liao, Vaughan, 2023; Unesco, 2024; Velez, Rister, 2024).

Tais entidades devem promover debates e encontros regulares com especialistas interdisciplinares e intersetoriais para analisar, de forma contínua, as implicações de longo prazo da IA Generativa em áreas acadêmicas, como a produção de conhecimento, direitos autorais, reformulação de currículos e avaliação, além da colaboração humana e dos impactos nas dinâmicas sociais.

É fundamental fomentar uma cultura de transparência e responsabilidade, na qual os pesquisadores são encorajados a compartilhar suas experiências e desafios no uso de IA. Ao abordar esses desafios éticos de forma proativa e reflexiva, a comunidade acadêmica pode qualificar o debate e fomentar usos que maximizem as benesses dessas ferramentas e mitiguem ao máximo seus malefícios (Chetwynd,2024; Resnik,Hosseini, 2024; Stahl, Eke, 2024).

VIII - Letramento em IA para Pesquisadores

Antes da apresentação de questões práticas, ressaltamos a ideia de que o letramento em Inteligência Artificial emerge como elemento relevante na formação contemporânea de pesquisadores, estabelecendo-se como via essencial para garantir o protagonismo humano no desenvolvimento científico (Limongi, Marcolin, 2024; Unesco, 2024; União Europeia, 2024). A integração responsável da IA no ambiente acadêmico demanda o desenvolvimento de competências específicas que permitam aos pesquisadores utilizarem estas ferramentas de maneira crítica e eficaz, sem comprometer a integridade e a qualidade da produção científica (Dabis, Csáki, 2024; Peres, 2024). Em busca de uma conceituação teórica para definir, ensinar e avaliar o letramento em IA, são propostas quatro dimensões, a saber: conhecer e compreender, usar e aplicar, avaliar e criar, e abordagem de questões éticas (Ng et al., 2021), como proposto por este guia.

Este processo formativo fundamenta-se em competências essenciais que se complementam e se reforçam mutuamente. O conhecimento técnico proporciona a compreensão dos princípios básicos de IA, incluindo seus algoritmos e arquiteturas, enquanto as habilidades práticas permitem a aplicação efetiva dessas ferramentas em pesquisas concretas. Já o pensamento crítico e a postura ética possibilitam a avaliação das limitações, vieses e implicações sociais da tecnologia. Finalmente, a compreensão contextual permite adaptar o uso da IA a diferentes campos de pesquisa, e a capacidade de colaboração humano-IA otimiza a integração entre recursos tecnológicos e expertise humana (Ray, 2023).

Entretanto, qualquer iniciativa de letramento em IA nas instituições de pesquisa requer uma abordagem sistemática que começa com a disponibilização de infraestrutura adequada e acesso equitativo a recursos computacionais. A capacitação docente torna-se fundamental através de programas de treinamento continuado e de workshops práticos, permitindo que os professores integrem efetivamente a IA em suas práticas pedagógicas e de pesquisa. Logo deve-se planejar desde o início uma integração curricular, mesclando atividades práticas supervisionadas e projetos interdisciplinares, permitindo aos pesquisadores em formação desenvolverem suas habilidades progressivamente (Velez, Rister, 2023).

O desenvolvimento destas competências deve focar na formação de pesquisadores capazes de analisar criticamente resultados gerados por IA, identificar limitações e vieses, e validar informações de maneira rigorosa. O uso ético e responsável das ferramentas de IA, considerando suas implicações sociais e questões de privacidade, torna-se parte integrante do processo formativo. A aplicação prática dessas habilidades materializa-se na integração da IA em projetos de pesquisa, no desenvolvimento de metodologias híbridas e na garantia de reprodutibilidade dos resultados.

Esta abordagem garante que os pesquisadores mantenham seu protagonismo no processo científico, utilizando a IA como ferramenta de apoio sem comprometer sua capacidade crítica e criativa, fomentando que a integração da IA no ambiente acadêmico fortaleça, em vez de diminuir, o papel central do pesquisador na produção do conhecimento (Dwivedi *et al.*, 2023; Limongi e Marcolin, 2024).

Princípios práticos

Entendida como um processo sistemático, amplo e complexo de compreensão do mundo, a ciência e seu método podem ser afetados em suas diversas etapas pelas ferramentas de Inteligência Artificial Generativa (Ramos, 2023; Sohail *et al.*, 2023; Bail, 2024; Cong-Lem *et al*., 2024; Hassan *et al*., 2024; Khan *et al*., 2024; Sampaio *et al*., 2024a), porém o processo em si não deve ser substituído. Entre outras coisas, isso significa compreender que a ciência não se limita a seus produtos (artigos, teses, dissertações etc.) e que a pesquisa nunca deve ser baseada em "apertar botões".

Diante disso, em um momento em que há dúvidas e incertezas, e considerando que grande parte dos guias sobre usos da IAG se focam apenas em princípios normativos amplos, como transparência e integridade científica, optamos por explorar seus usos cotidianos, apesar de nosso objetivo não ser esgotar todas as fases do processo acadêmico. Além de reflexões iniciais sobre as questões éticas associadas, cada uso é acompanhado de cuidados que o pesquisador deve tomar.

1 – Exploração inicial de ideias

Os Grandes Modelos de Linguagem foram treinados em extenso conhecimento humano disponível na Internet, sendo que um de seus limites conhecidos é a reprodução do conhecimento existente, sem a capacidade de serem realmente criativos (Santaella, 2023). Esta limitação não significa, entretanto, que não possam funcionar para o teste de ideias e mesmo para a elaboração de hipóteses (Almeidas, Nas, 2024; Cong-Lem *et al*., 2024; Khan *et al*., 2024).

Parece-nos razoável, então, que esta tecnologia possa ser útil para a exploração inicial de ideias, atividade relacionada ao planejamento de pesquisa. Atualmente, encontramos *prompts* que colocam a IA em papeis acadêmicos, como por exemplo, mentores, orientadores, revisores de periódicos, membros de comitês de seleção ou até "debatedores socráticos". Assim, a IAG pode ser utilizada para ajudar um pesquisador a verificar pontos fracos e fortes de sua pesquisa, assim como refletir sobre como melhorá-la de forma geral (Lopezoza, 2023; Dabis, Csáki, 2024; Unesco, 2024).

Particularmente, a IAG, por meio de suas características probabilísticas e combinatórias, pode gerar ideias a respeito de um determinado tema. No âmbito da pesquisa científica, ela se reflete na "chuva de ideias" (*brainstorming*) para a sugestão de temas, problemas, objetivos e hipóteses de pesquisa. Quando utilizada de forma iterativa, em sucessivos passos que replicam as etapas da definição de um projeto de pesquisa, podem constituir uma ferramenta auxiliar para o pesquisador. Entretanto, ressaltamos que o domínio da metodologia de pesquisa, circunscrita às práticas e costumes de sua área de conhecimento, é condição para o efetivo aproveitamento dessa ferramenta. Além disso, o conhecimento conceitual e do estado da arte da pesquisa na área também são necessários para uma avaliação crítica das ideias sugeridas.

Para pesquisadores em formação, particularmente, as capacidades da IAG se apresentam como faca de dois gumes. Por um lado representam um campo de prática para o exercício do planejamento de pesquisa. Por outro, a facilidade de uso e aparente qualidade dos resultados, pode levar a uma aceitação leviana das respostas e a uma perda do pensamento crítico, podendo chegar à excessiva dependência desses sistemas.

⚠️ Cuidados

Assim como em outros contextos, a qualidade dos resultados dos testes de ideias com auxílio da IAG depende dos dados de treinamento e da forma como os *prompts* são executados. Quanto mais complexa e específica a temática, maiores as chances das IAs gerarem apenas senso comum ou devolverem generalizações e superficialidades.

No atual momento, as IAs Generativas foram significativamente treinadas com maior intensidade em língua inglesa. Assim, as ideias geradas ou analisadas frequentemente seguirão uma visão anglo-saxã de ciência e de acordo com os cânones de determinada área (viés cognitivo). Da mesma maneira, *elas podem utilizar conteúdo protegido por direitos autorais, de forma que é necessário bastante cuidado para não incorrer em plágio inadvertido.* Nunca deixe de verificar se as respostas de uma interação são, de fato, inéditas. Em caso negativo, busque e cite os autores originais.

Por outro lado, em atividades complexas que envolvem múltiplas áreas do cérebro humano, como criatividade e inovação, a IA ainda não é capaz de

gerar produtos originais, de forma independente. Por ser treinada em trabalhos publicados, a IA pode sugerir questões de pesquisa já conhecidas, em vez de temas de ponta, e pode priorizar variáveis, métodos e descobertas de subdisciplinas dominantes nos dados de treinamento (Sampaio *et al.*, 2024a; Unesco, 2024; União Europeia, 2024).

Finalmente, as IAG foram basicamente treinadas para ser ferramentas amigáveis e "educadas" (Santaella, 2023), que frequentemente irão se desculpar por eventuais erros ou lacunas. Na prática, isso quer dizer que as IAG frequentemente buscam inicialmente ser agradáveis aos seus usuários somente discordando se devidamente solicitadas. Uma parte significativa da pesquisa, da geração de novas ideias, conceitos e teorias está diretamente relacionada com a discordância, com o tensionamento de ideias, algo não realizado geralmente pela IAG[33]. Isso em certos casos então vai demandar *prompts* adequados e, em outros, apenas reconhecer que a IAG não deve sempre ser a melhor ferramenta para este caso.

Uma orientação prática seria a de formular e rascunhar ideias de forma tradicional, antes de iniciar uma conversação com o chat de IA. Esta sutil mudança de atitude coloca a tecnologia como um "par" (*peer*) que irá colaborar na discussão e avaliação das ideias em questão. A experiência com sistemas como o ChatGPT ressalta que a inteligência humana continua essencial para a formulação de perguntas interessantes, geração de conhecimento novo e para a condução de pesquisas, de forma geral (Block, Kuckertz, 2024). Seguindo a postura que adotamos neste guia, a IA pode servir como um auxílio, não como substituto do raciocínio, cognição e criatividade humanos.

Aqui, fazemos um paralelo com um interessante livro de Renato Gonçalves (2023), que brinca com a ideia de "cr(IA)ção", ou seja, de novos processos criativos que vão ser mediados e potencializados pela Inteligência Artificial, mas nunca substituídos. O livro trata sobre criatividade verbal, visual e sonora e mostrando usos ativos e criativos das IAG para alcançar resultados mais interessantes. O livro termina com sete dicas de cr(ia)ção que nos parecem interessantes de serem reproduzidas pelo meio acadêmico:

- A Inteligência Artificial não faz nada sozinha. Convide-a para um diálogo criativo.
- Não seja um apertador de botão. Descubra sempre a sua intencionalidade criativa.
- Sem repertório criativo e crítico, você estará sempre na superfície da Inteligência Artificial.
- O jogo faz parte do processo criativo e a Inteligência Artificial é um baralho de possibilidades.
- Nem tudo se resume à Inteligência Artificial. Ela pode estar apenas em uma parte do seu processo criativo tradicional.
- A Inteligência Artificial não substitui ninguém, mas pode facilitar algumas etapas.
- Tire proveito do melhor da inteligência artificial, tire proveito do melhor da inteligência humana.

Apesar de ser frequentemente vista como um processo técnico e rígido, a pesquisa acadêmica definitivamente é enriquecida quando há criatividade envolvida na sua concepção.

2 – Busca de materiais acadêmicos

As ferramentas de IA oferecem uma capacidade notável de processar rapidamente grandes volumes de publicações, identificando as mais relevantes para uma área de estudo específica, além de fornecer resumos precisos e interações entre os trabalhos. Limongi (2024) ressalta que essa capacidade não apenas economiza tempo, mas também assegura que os pesquisadores estejam cientes dos desenvolvimentos mais recentes e pertinentes a seus respectivos campos. Além disso, a IA pode detectar tendências emergentes e lacunas na literatura, direcionando os pesquisadores para áreas inexploradas que podem ser frutíferas para investigação.

Várias plataformas digitais baseadas em IA permitem fazer a busca de materiais acadêmicos por meio de perguntas, sendo algumas de uso geral, como Copilot, Gemini e *Perplexity*, *You.com*, enquanto outras são exclusivamente baseadas em fontes acadêmicas, como *Consensus*, *Elicit*, *Keenious*, *Iris*, *SciSpace*, *Scite*, *Undermind* etc. Sempre prefira as acadêmicas. É preciso atentar para a escrita e/ou utilização de *prompts* que capturem a intencionalidade da pesquisa e para o fato de que pequenas diferenças em sua formulação podem levar a resultados significativamente distintos (Schulhoff *et al*., 2024). Tais sistemas são particularmente interessantes quando você precisa de materiais acadêmicos sobre uma questão bastante específica e têm dificuldades de encontrar por meios tradicionais. Aqui o caminho

33. Agradecemos a Ricardo Fabrino por esta reflexão.

é se aproveitar das vantagens da busca semântica (sentido das frases) sobre as buscas mais léxicas (por exemplo, palavras-chave)[34]. Portanto, é interessante fazer diversos testes antes de uma efetiva coleta.

Outras plataformas permitem a visualização de redes de citação de artigos, a exemplo de *Connected Papers*, *Litmaps*, *Nested Knowledge*, *Research Rabbit*, geralmente funcionam por meio da inserção de referências para alimentar o início da rede. Essa inserção pode se dar por título do artigo, DOI ou mesmo por arquivos de referências (.bib e .ris). Para além das redes, a principal vantagem dessas ferramentas reside na identificação de referências tanto a estudos antecedentes quanto a posteriores, compondo uma rede de citações potencialmente mais abrangente e facilitando encontrar referências diretamente conectadas ao estudo em questão. No entanto, é fundamental considerar que o ponto de partida da rede de citação já reflete um viés inerente ao pesquisador e que isto pode influenciar diretamente nos resultados obtidos. Provavelmente, o mais lógico seria usar essas plataformas na forma de uma busca de citações (citações em cascata, ver Wohlin, 2014), melhor indicada para pesquisas em temáticas novas, pouco abordadas e afins.

⚠ Cuidados

Diante do atual estado da tecnologia, revisões de literatura conduzidas por IAG podem ser problemáticas, pois o resultado pode carecer da profundidade acadêmica esperada. Determinadas comunidades de pesquisa demandam uma cobertura de conteúdo científico, abrangência de citações e fundamentação teórica mais robustas, o que pode não ser alcançado através da automatização. Além disso, diversas IAG, especialmente as não acadêmicas, podem gerar citações incorretas ou vincular o autor a fontes inadequadas (Bail, 2024; Jesus, Santarém Segundo, 2024; Rodrigues *et al.*, 2024).

Em todos os casos, é importante atentar-se às fontes que tais plataformas acessam. A grande maioria funciona sob a API da base indexadora *Semantic Scholar*, que é uma base ampla, mas que certamente não contém algumas das principais editoras e periódicos de cada área, muitas das quais estão presentes apenas em bases acadêmicas mais tradicionais, a exemplo de Pubmed, SciELO, Scopus[35], Web of Science e afins. Então, tais ferramentas são bastante limitadas para revisões sistemáticas de literatura se não houver um cruzamento com outras bases de dados. Além disso, chatbots, como ChatGPT[36], Claude, Mistral e Maritalk não devem ser usados como fontes primárias de pesquisa e de dados, pois muitas vezes esses modelos não são conectados à internet e podem dar informações desatualizadas ou mesmo incorretas.

Outro aspecto a se observar é a capacidade ou não de tais soluções coletarem artigos publicados em português e/ou em revistas brasileiras. Em nossos testes, algumas eram mais capazes de encontrar este tipo de material, em relação a outras. Assim como acontece na escrita, tais plataformas foram geralmente pensadas para a língua inglesa e funcionam melhor com *prompts* redigidos nela.

Block, Kuckertz (2024) discutem em especial como a inteligência artificial afeta as revisões sistemáticas de literatura (RSLs) realizadas por humanos. A IA já automatiza etapas como busca e síntese de dados, mas os autores argumentam que o papel humano ainda é necessário em ao menos três momentos: formulação de perguntas, seleção de estudos e interpretação crítica. Uma RSL eficaz vai além da mera síntese de estudos e busca interpretar criticamente os achados, propondo novas direções de pesquisa. A IA baseia-se em dados históricos, o que pode limitar a sua capacidade de antecipar tendências ou desafios emergentes, enquanto os pesquisadores humanos são capazes de problematizar e teorizar com base em suas interpretações individuais. Os autores inclusive valorizam a serendipidade, ou o acaso nas descobertas, que enriquece a análise ao introduzir elementos não planejados, algo que só pode vir da interação humana.

34. Apesar de não se tratar de uma efetiva pesquisa, este relatório apresenta alguns dados relevantes sobre as vantagens de pesquisas semânticas. TAY, Aaron. Academic search and discovery tools in the age of AI and large language models: An overview of the space. (2024). **AI for Research Week.** Disponível em: https://ink.library.smu.edu.sg/ai_research_week/Programme/Programme/1. Acesso 5 nov. 2024.

35. Durante a escrita deste material, já estava disponível a ferramenta de IAG da Scopus, que pode ser uma alternativa interessante a tais buscadores acadêmicos, porém a mesma ainda não estava disponível pelo periódico CAPES. Disponível em: https://www.elsevier.com/pt-br/products/scopus/scopus-ai. Acesso em 2 nov. 2024.

36. No momento da escrita deste material, o ChatGPT acabara de lançar uma versão capaz de buscar a internet e que era capaz de identificar literatura acadêmica, porém nossos testes iniciais evidenciaram não ser comparável a fontes exclusivamente acadêmicas. Mais a respeito do buscador em: https://istoe.com.br/istoegeral/2024/11/03/o-chatgpt-vai-destronar-o-google-conheca-a-nova-ferramenta-de-busca/. Acesso 3 nov. 2024.

Portanto, embora as ferramentas de IA ofereçam vantagens significativas na busca e síntese de materiais acadêmicos, elas devem ser utilizadas como um complemento, e não como substituto do julgamento crítico e da expertise do pesquisador. A verificação cuidadosa das fontes, a avaliação da profundidade e relevância do conteúdo gerado, e a complementação com pesquisas em bases de dados tradicionais continuam sendo práticas essenciais para garantir a qualidade e rigor acadêmico do trabalho de revisão de literatura. Para algumas aplicações e testes iniciais deste tipo de uso, conferir Jesus, Santarém Segundo (2024), Rodrigues *et al.* (2024) e Sampaio *et al*. (2024b). [

3 – Leitura e resumo de materiais acadêmicos

Ferramentas de IAG e processamento de linguagem natural são particularmente eficazes na verificação de textos escritos por humanos da mesma forma que são para gerar tais textos. Com isto, um uso particular se torna atrativo para a comunidade da pesquisa científica: a leitura de material bibliográfico, tendo em conta o grande volume da produção de qualquer área do conhecimento. Seja para na elaboração de uma revisão bibliográfica, ou simplesmente para que o pesquisador possa se manter atualizado em sua especialidade, o volume crescente de leitura é um desafio.

Portanto, várias soluções de IA permitem que você tenha "diálogos" com esses arquivos, podendo solicitar para encontrar determinadas informações no texto, resumir determinados pontos ou mesmo explicar partes específicas. Novamente, temos leitores gerais, como os Grandes Modelos de Linguagem (ChatGPT, Claude, *Notebook LM* da Google etc.), ferramentas que permitem o envio dos arquivos de texto, a exemplo de PDFs originais para o início da análise, como *ChatDoc*, *ChatPDF*, *Humata*, *My Reader*, etc. e outras com foco acadêmico, como *Avid Note*, *Elicit*, *Explain Paper*, *SciSpace* e *Research Rabbit*. Entre estas, algumas permitem o "diálogo" com múltiplos artigos, facilitando buscar semelhanças e diferenças entre os materiais, além de estabelecer redes conceituais. Novamente, 1) os melhores resultados geralmente são vistos naqueles dedicados a materiais acadêmicos, especialmente quando se trata de artigos científicos; e 2) os resultados dependem de bons *prompts*.

Outras ferramentas são especializadas no resumo de materiais, inclusive de natureza acadêmica, como *Scholarcy* e *SciSummary*. São aplicativos particularmente interessantes para se ter sumários práticos de vários trechos dos artigos. Geralmente, tais ferramentas trazem algumas vantagens como exportação de referências, gráficos e figuras dos artigos.

Tais possibilidades não se restringem à leitura de materiais no formato de texto. No atual momento, inúmeras soluções baseadas em IAG já são capazes de, rapidamente, resumir áudios e vídeos online (e.g. YouTube), como as extensões *Glasp*, *Merlin* e *Harpa*. Também começam a surgir experimentações interessantes, na quais as ferramentas reorganizam as ideias centrais de textos ou vídeos em outros formatos, como, por exemplo, na forma de um podcast, no qual dois personagens debatem sobre o assunto presente no conteúdo original, já disponível no Notebook LM da Google[37].

Estas ferramentas parecem particularmente interessantes para dois aspectos principais: 1) a decisão de ler ou não adequadamente o material após essa exploração inicial; 2) fazer resumos automatizados de uma grande quantidade de referências para pesquisas específicas. No segundo caso, pode haver interesse em identificar métodos ou conceitos compartilhados por diferentes estudos, por exemplo.

⚠️ Cuidados

De forma geral, *a IA para leitura de materiais acadêmicos deve ser usada com cautela, considerando seus limites e questões de segurança.* Evite subir (*upload*) materiais que sejam inéditos e/ou sensíveis, pois estes poderão ser usados para o treinamento dos modelos de linguagem. Da mesma forma, é preciso precaução com as questões de direitos autorais dos materiais a serem utilizados. Evite, em especial, fazer o *upload* de materiais ainda não publicados. Adote como prática ativar as configurações que previnem o uso dos dados para treinamento do modelo.

Por sua vez, os "diálogos" e resumos gerados por IAs estão fortemente suscetíveis aos limites e *bias* dos modelos no quais foram treinados, podendo apresentar informações falsas, incompletas, inventadas, incorretas ou mesmo baseadas no

37. ISMAIL, Youssef. Illuminate: nova ferramenta do Google transforma artigos científicos em podcasts. **PodNotícias**, 26 jul. 2024. Disponível em: https://podnoticias.com.br/illuminate-nova-ferramenta-do-google-transforma-artigos-cientificos-em-podcasts/. Acesso em: 30 set. 2024.

treinamento do modelo e não nos arquivos fornecidos pelo pesquisador. Ou seja, sempre é possível que aconteça o processo de alucinação ou simplesmente que a máquina não gere um resultado compatível com o conteúdo original.

O princípio da agência humana permanece central no processo de leitura acadêmica. Resumos, quer sejam gerados por IA ou não, não substituem a leitura completa e crítica de materiais, com anotações e reflexões. Dwivedi et al. (2023, p. 25) alertam para o risco de estudantes, já afetados por menor capacidade de atenção e redução na leitura de livros, entrarem em um "modo letárgico" com o uso excessivo de resumos e IA. Portanto, manter o equilíbrio entre o uso de tecnologias de apoio e o engajamento direto com materiais acadêmicos é importante para o desenvolvimento intelectual e a capacidade de pensamento crítico.

O processo de leitura, embora muitas vezes desafiador, é fundamental para o desenvolvimento de um pesquisador, proporcionando ganhos cognitivos que podem ser comprometidos pela dependência excessiva de ferramentas de IA. Como destacado pela Unesco (2024, p. 38), as habilidades fundamentais de alfabetização e letramento científico continuarão essenciais no futuro. Há, de fato, um potencial das IAG modificarem a forma como se dá o aprendizado; contudo, como sempre, ela será útil enquanto for apenas uma ferramenta para auxiliar o pesquisador.

4 – Escrita

Com apelo sedutor para uma tarefa complexa e trabalhosa, o uso de ferramentas, plataformas e soluções de Inteligência Artificial Generativa para a escrita é certamente o ponto mais complexo, polêmico e aberto a debates[38]. Um número razoável de periódicos e editoras acadêmicas, assim como universidades pelo mundo admitem, até o momento, apenas o uso de tais instrumentos para a correção em níveis mais técnicos da escrita, a exemplo da correção de gramática e ortografia (Elsevier, 2023b; Sage, 2023). Se usado apenas para correções desta natureza, não há necessidade de reportar o uso de tais ferramentas, assim como não é usual relatar que fez uso do corretor de processadores de texto (a exemplo do Microsoft Word) ou de uma ferramenta como o *Grammarly*, por exemplo. Caso o uso seja mais aprofundado, deve-se adotar os critérios de transparência do uso de tais ferramentas. Verifique as diretrizes da editora, associação científica ou periódico a ser submetido antes de passar deste ponto.

Em outras visões, o uso de tais modelos generativos para agregar, resumir, expandir, parafrasear e alterações mais básicas em termos do texto, é aceitável. Estas ferramentas poderiam auxiliar em questões de gramática, pontuação, complexidade e mesmo voz, no sentido de autoria própria, para jovens estudantes e pesquisadores. Podendo, inclusive, facilitar o trabalho em equipe e permitindo que autores incrementem a qualidade do texto a ser compartilhado com colaboradores, professores e orientadores, podendo ocasionar um efeito positivo no senso de eficácia de jovens escritores (Boyd-Graber; Okazaki; Rogers, 2023; Susarla *et al*., 2023).

Algumas universidades americanas também aceitam o uso para determinadas tarefas, tais "como redação de abstracts, análise de grandes volumes de dados e formatação de documentos para periódicos específicos – tarefa que tem baixíssimo impacto na pesquisa em si, mas que demandam muito tempo do pesquisador" (Soares *et al*., 2023, p. 83). Há algumas sugestões sobre a utilização em sessões mais padronizadas e rígidas, como *métodos,* e a descrição pura de *resultados* (Korinkek, 2023), além do uso para geração ou melhoria de *título, resumo* e *palavras-chave*. **Não obstante, no atual momento, tais empregos estão longe de ser um consenso**.

Aqui, a grande preocupação é sempre se ater a possíveis distorções do modelo e a conferência sobre a inserção de algum erro ou inexatidão no conteúdo gerado. Em outras palavras, no atual momento, mesmo que o *prompt* de entrada seja apenas de revisar um trecho, ele poderá incluir informações novas ou mesmo retirar outras[39]. Então, é muito importante que sempre haja humanos no processo (*human in the loop*) e os autores estarem cientes de que vão assumir o conteúdo final integralmente (Sage, 2023).

⚠ Cuidados

Em hipótese alguma, deve-se apenas "copiar e colar" o texto gerado por uma IA em quaisquer tipos

38. O impacto crescente da Inteligência Artificial na geração de linguagem levanta questões sobre originalidade, autoria, influência, verdade, conhecimento, liberdade e a relação entre o humano e a máquina. Entretanto, os "robôs escritores" possuem raízes históricas antigas (Alexander et al. 2021).

39. Isso inclusive pode acontecer quando se usa a IAG para acertar as citações de documentos (i.e. ABNT, APA, Vancouver etc.) e em outras tarefas rotineiras similares.

de materiais acadêmicos ou avaliações. Da mesma forma, deve-se ter cautela nos usos de ferramentas de IA escritoras que (a) aumentam o tamanho do texto; (b) parafraseiam o escrito; (c) mudam o tom da escrita; principalmente aquelas que (d) escrevem ou continuam a escrever o texto; e, mesmo, (e) citam referências acadêmicas, como *Jenni*, *Paperpal*, *SciSpace*, *Smoodin*, *Trinka* etc.[40]. Já há extensões de navegadores, como *Compose Ai* e *Text Cortex*, que são capazes de editar documentos online, a exemplo do Google Docs. Simultaneamente, a Google trabalha na sua própria ferramenta para seus serviços online, enquanto a Microsoft implementa o Copilot dentro do pacote Office, que já permite a geração e correção de texto no Word[41]. Todas essas opções só podem ser usadas com extrema cautela se houver regras que as permitam no local de sua produção (por exemplo, regramentos e normas universitárias) ou submissão (como políticas editoriais de periódicos, congressos e editoras científicas).

Como já dito em outros momentos, as IAs estão frequentemente buscando materiais para o treinamento de seus modelos, o que pode incluir o processo de escrita. Novamente, atente-se com dados inéditos ou sensíveis, pois eles poderão ser usados por tais soluções para seus negócios. *Evite produzir o texto inteiro nestas soluções se não houver clareza sobre esse uso ou não. Lembre-se ainda que o texto produzido por estas ferramentas pode ter direitos autorais e/ou ser baseado diretamente nas ideias de outras pessoas, podendo se tratar de plágio inadvertido* (Cotton *et al*., 2023; Barreto, Ávila, 2023). Além disso, inclua apenas as citações que tenham, de fato, sido lidas e consideradas relevantes para o trecho sendo escrito. **Nunca use as citações sugeridas por IAs acadêmicas sem checar os materiais originais**.

No momento da produção deste material, já se notava que os grandes modelos de linguagem também apresentam seus próprios vícios de linguagem. Um estudo sobre produção acadêmica em inglês, por exemplo, demonstrou o uso excessivo da palavra "delve", que é geralmente usada pelo ChatGPT, mas incomum no cotidiano dos pesquisadores (Kobak *et al.*, 2024)[42]. As IAGs enfrentam limitações significativas na produção de textos inéditos de qualidade, geralmente não tendo coesão, o que compromete a fluidez e a conexão lógica entre as ideias. *Portanto, é vital a revisão humana final do texto para que os vícios de linguagem do modelo não se tornem os seus***.**

Além disso, a abordagem de questões teórico-metodológicas tende a ser superficial, sem a profundidade necessária para análises complexas. Outro problema recorrente é a textualização vaga, caracterizada pelo excesso de adjetivos e advérbios, o que dilui a precisão argumentativa. A ausência de exemplificação também enfraquece o texto, tornando-o menos convincente e difícil de sustentar em discussões acadêmicas ou científicas (Cotton *et al*., 2023; Chen *et al.*, 2024; Chetwynd, 2024; Mangold, Ream, 2024; Peres, 2024). Lindebaum e Fleming (2024) chegam a afirmar que a reflexão humana permite uma reinterpretação constante da realidade, algo que a IA não consegue realizar. A "gramática transcendental" do pensamento humano permite a criação de significado além dos dados, o que é essencial para descobertas acadêmicas. A IA tende a homogeneizar o conhecimento, o que, segundo os autores, prejudica a evolução da gestão responsável, que depende de interpretações diversas e inovadoras.

Na mesma linha, é importante destacar que a escrita acadêmica é uma das principais habilidades a serem adquiridas por pesquisadores e que diversos ganhos de aprendizado estão diretamente conectados ao ato da escrita. O desenvolvimento de uma autoria e de um estilo único também estão conectados a uma prática constante de tal habilidade (Cruz, 2020). Concordamos que "a linguagem é uma parte importante do que nos torna humanos. [...] A linguagem escrita constitui grande parte de nossa sociedade, nossas regras, normas, expectativas e

40. O objetivo aqui não é indicar a fórmula para trapacear, mas alertar professores e pesquisadores sobre a fácil disponibilidade de tais ferramentas.

41. Microsoft. "Copilot no Word: Explorar como você pode usar o poder da IA no Word". Disponível em: https://copilot.cloud.microsoft/pt-br/copilot-word. Acesso 5 nov. 2024.

42. Estas palavras certamente poderão ser diferentes no momento da leitura, porém, no momento da escrita, já notamos que as IAG, em português, tendem a abusar de adjetivos para caracterizar a relevância da pesquisa, a exemplo de "crucial", "eficaz", "essencial", "fundamental", "imperativo", "inspirador", "revolucionário", "significativo", "valioso", "único", "vital" e de verbos relacionados à pesquisa, como "aprimorar", "contribuir", "empregar", "evitar", "implementar", "perpetuar", "promover", "reduzir", "refletir", "resultar", "sublinhar", "utilizar". Já se verifica também o uso de expressões não usuais como "multifacetado" e "intrincado". Curiosamente, como não tem uma tradução direta para o português, é normal vermos um uso excessivo também de "insight". Este ponto certamente merece aprofundamento e pesquisas específicas para o português brasileiro. Uma primeira aproximação: XAVIER, Mateus. "Analisando as Palavras entre Textos Originais e Textos Gerados pelo ChatGPT"*. **Medium**, 24 abr. 2024. Disponível em: https://medium.com/@mateus.xavier/analisando-as-palavras-entre-textos-originais-e-textos-gerados-pelo-chatgpt-117e334b8532 . Acesso em: 2 nov. 2024.

rotinas" (Dwivedi et al., 2023, p. 39). Neste sentido, o uso das IAs deve ser acompanhado de prudência de modo a evitar impactos negativos em tais processos da construção da autoria.

O estudo de Alvero *et al*. (2024), por exemplo, comparou o estilo de escrita de 150 mil cartas de admissão para universidades públicas dos Estados Unidos com mais de 25.000 redações geradas pelos GPT 3.5 e GPT 4. Os resultados evidenciam que as redações geradas por IA se aproximam àquelas escritas por homens de classes sociais mais elevadas, evidenciando tanto o viés da ferramenta quanto a diminuição da diversidade de perspectivas na escrita.

Por estas razões, como possibilidade prática, recomendamos sempre incluir nos *prompts* (ver Apêndice 2) pedidos da explicação, por parte do chatbot, das razões pelas quais ele revisou um texto de determinada forma ou sugerir duas ou três correções alternativas ao texto original. No caso de ferramentas que já incluem a reescrita com alterações, uma boa prática é nunca usar diretamente o texto revisado, mas colar abaixo do seu e verificar manualmente as diferenças entre as duas ou mais versões. A partir da comparação, uma terceira versão pode ser elaborada, manualmente, com a seleção dos melhores trechos da versão original e da revisada. É importante que este processo sempre tenha uma lógica de aprendizado e nunca de dependência.

5 – Análise e apresentação de resultados

Grandes Modelos de Linguagem, como ChatGPT e Claude, geralmente incluem funções específicas para análise de dados, usando servidores externos de Python ou similares para a tarefa (Hassan *et al*., 2023; Ray, 2023; Sohail *et al*., 2023; Lehr *et al*., 2024), sendo úteis em diversos tipos de análises. Da mesma maneira, essas e outras IAGs podem facilmente produzir quadros, tabelas e visualizações de dados, como grafos e gráficos, algumas a exemplo de *Heuristi.ca* e *Napkin* são capazes de transformar praticamente qualquer material escrito em visual. Também é possível realizar análises automatizadas em larga escala, diretamente em modelos de fronteira ou por meio de API, particularmente úteis para análises de dados qualitativos (Hamilton *et al*., 2023; Sampaio *et al*., 2024b).

Temos aqui uma situação nebulosa e menos consensual. Por um lado, a maioria das diretrizes consultadas para a construção deste texto proíbe ou não recomenda o uso de tais tecnologias para a análise de dados. Os diversos problemas já relatados neste guia aqui se acumulam. Por conta dos vieses e alucinações, tais modelos podem inventar ou falsear resultados, eventualmente falhando na condução de uma pesquisa qualificada. Por exemplo, foram identificados momentos nos quais a resposta não é a esperada[43], prejudicando a confiabilidade e validade dos dados.

Encontramos também várias críticas à falta de transparência e de replicabilidade dessas ferramentas de IA; o cerne desses modelos são autênticas caixas-pretas, protegidas por segredos industriais. Há poucas informações públicas sobre o seu funcionamento e atualizações, o que afeta os resultados. Por exemplo, no momento que este guia está sendo escrito, um dos modelos ainda disponíveis pela OpenAI é o ChatGPT-4o-preview, significativamente diferente da sua primeira versão. No entanto, ainda assim não dispomos de detalhes sobre essas mudanças (se, por exemplo, trata-se de uma versão 4.1 ou similar). Diferentemente de *softwares* acadêmicos estáveis e rastreáveis, LLMs são periodicamente atualizados e as versões anteriores não são preservadas, impedindo a replicabilidade (Liao, Vaughan, 2023; Sampaio *et al.*, 2024a). Retomando o exemplo, a capacidade de análise da atual versão do ChatGPT 4 é muito superior em desempenho de sua versão de lançamento. Portanto, muitos aspectos estão fora do controle do pesquisador, dificultando quesitos essenciais da pesquisa científica, tais como são a transparência, confiabilidade, replicabilidade e validação dos resultados.

Por outro lado, pesquisadores de todo o mundo têm testado a capacidade dos modelos de IAG para diversos tipos de análise científica. Existem testes para verificar quão bom eles são enquanto cientistas de dados (Hassan *et al*., 2023; Lehr *et al*., 2024), como partes de experimentos (Bubeck *et al*., 2024), como ferramentas de análise textual automatizada (Bail, 2024; Ziems *et al*., 2024), incluindo análise de sentimentos (Silva, Serrano, 2023), como analistas de dados qualitativos de forma ampla (Hamilton *et al*., 2023; Sampaio *et al*., 2024b)[44], entre outros usos (Sohail *et al*., 2024).

43. ALBUQUERQUE, Karoline. OpenAI confirma que ChatGPT está mais 'preguiçoso'. **TecMundo**, 13 dez. 2023. Disponível em: https://www.tecmundo.com.br/software/274936-openai-confirma-chatgpt-preguicoso.htm. Acesso em: 26 set. 2024.

44. A Revista Pesquisa Qualitativa organizou um interessante dossiê trazendo discussões e aplicações de IAG em pesquisas qualitativas. O dossiê pode ser acessado na íntegra em: https://editora.sepq.org.br/rpq/issue/view/31.

⚠ Cuidados

Mantenha uma postura cética em relação aos resultados gerados por Inteligência Artificial Generativa. Adote uma abordagem científica, experimentando diferentes *prompts*[45], testando em outros *softwares*, utilizando outras linguagens de programação ou mesmo manualmente (por exemplo em análises qualitativas). Experimente com parte dos resultados, a exemplo de adotar uma amostra do conjunto total de dados e testar os resultados por outras formas. Documente os resultados encontrados, incluindo possíveis divergências e os divulgue junto a sua análise (Liao, Vaughan, 2023).

Atente-se a dados sensíveis, inéditos ou que impliquem cuidados de segurança. Por exemplo, ao pedir que um modelo de linguagem avalie entrevistas qualitativas em profundidade, tal conteúdo poderá ser usado para o treinamento do modelo, quebrando princípios éticos de privacidade, de anonimato e de segurança dos dados dos participantes. Essa recomendação vale para levantamentos (*surveys*), testes clínicos, grupos focais, entre outros. Lembre-se: dados inéditos poderão também ser usados para o treinamento dos modelos.

Priorize análises que sejam plenamente replicáveis, como, por exemplo, por meio da utilização de linguagens de programação, como R e Python, ou mesmo de modelos abertos de IAG (Bail, 2024). Disponibilize os *scripts*, assim como os *prompts* utilizados para gerar a análise. Seja explícito na seção de métodos do estudo sobre quais ferramentas de IA foram utilizadas, incluindo versão e data do uso. Se possível, compartilhe o *link* da interação com o modelo de linguagem.

Por questões de direitos autorais, imagens geradas por Inteligências Artificiais Generativas como o *Dall-e*, *Midjourney*, e o *Stable Diffusion*, não são aceitas pelas principais editoras, por exemplo, a Elsevier (2023b) e periódicos de renome, a exemplo da Nature (2023). Tampouco é permitido o uso de IA para alterar imagens em manuscritos, exceto ajustes básicos de brilho, contraste ou balanço de cores. Imagens geradas por IA só são permitidas se tal uso fizer parte do desenho de pesquisa ou dos métodos de pesquisa (como um experimento). Este uso deve ser descrito de forma reprodutível, na seção de métodos; repetindo: inclua detalhes do modelo de IA utilizado (nome, versão etc.), *prompts* e *link* para a conversa.

Enquanto isso, alguns *softwares* científicos estão começando a empregar modelos de linguagem em suas análises, com destaque para os CAQDAs (*Computer Assisted Qualitative Data Analysis*, programas para a análise de dados qualitativa assistida por computador) a exemplo de Atlas.ti[46], MAXQDA[47] e NVivo[48]. O fato de estarem incorporados ao *software* significa que fazem uso da API dos Modelos de Linguagem e que têm contratos que garantem a privacidade dos dados e o uso sem o treinamento de modelos[49]. Provavelmente, devem existir ajustes também para mitigar alucinações.

Independentemente, e considerando o momento atual, é importante checar as análises automatizadas realizadas por tais ferramentas, pois elas continuam incorporando distorções e vieses dos Modelos de Linguagem. Prefira sistemas especializados em análise de dados, como Data Squirrel, Julius, e Rows sempre conferindo suas informações de privacidade e segurança.

6 – Programação

Antes do lançamento do ChatGPT, vários ramos da ciência vinham progressivamente usando linguagens de programação, a exemplo de R, Python, Julia e Cobol para a análise de dados. Vimos nos últimos anos a rápida ascensão das profissões de cientistas de dados e analistas de dados, inclusive abrindo novos mercados laborais para acadêmicos. Apesar das curvas de aprendizado relativamente altas, estava se criando um ambiente de valorização de tal habilidade no mundo acadêmico (Frommherz, Langenhorst, 2022).

Novamente, as IAG impactam diretamente, pois além da linguagem humana, elas também foram treinadas em códigos de programação criados por

45. Algumas dicas da empresa que faz o MAXQDA para interagir melhor com IAGs para análise de dados qualitativos. Müller, Andres, Rädiker, Stefan. "'Chatting' With Your Data: 10 Prompts for Analyzing Interviews With MAXQDA's AI Assist" https://www.maxqda.com/blogpost/chatting-with-your-data-10-prompts-for-analyzing-interviews-with-maxqda-ai-assist. Acesso em 4 nov. 2024.

46. Disponível em: https://atlasti.com/research-hub/intentional-ai-coding. Acesso em: 14 out. 2024.

47. Disponível em: https://www.maxqda.com/help-mx24/ai-assist/what-is-ai-assist. Acesso em: 14 out. 2024.

48. Disponível em: https://lumivero.com/resources/blog/revolutionizing-text-data-analysis-with-ai-autocoding-with-nvivo/. Acesso em: 14 out. 2024.

49. Por exemplo, a especificação do Atlas.ti. Disponível em: https://atlasti.com/ai-data-security-and-privacy. Acesso em: 14 out. 2024.

humanos, apresentando funções para geração e correção. Portanto, elas podem incrementar rapidamente a capacidade de pesquisadores de gerar bons *scripts* para diferentes fases da pesquisa. Além disso, podem ser utilizadas para sugerir melhorias no código existente, identificar erros e otimizar o desempenho. Essas tarefas de depuração incluem a reescrita de funções para torná-las mais eficientes ou a recomendação de bibliotecas e pacotes adequados para tarefas específicas. Além disso, podem auxiliar na geração de documentação e comentários para o código, tornando-o mais compreensível e acessível para que outros pesquisadores possam trabalhar com os mesmos *scripts* no futuro (Bail, 2024; Sohail *et al*., 2024).

Entretanto, também cabe considerar que programadores iniciantes enfrentam obstáculos metacognitivos ao resolver problemas de programação, muitas vezes sem perceber como essas dificuldades limitam seu avanço. Enquanto a IA Generativa oferece suporte, um estudo recente encontrou que a ferramenta ampliou tais dificuldades, proporcionando uma ilusão de competência (Prather et al., 2024), o que exige mais pesquisas que possam subsidiar a aprendizagem e prática da programação.

⚠ Cuidados

Todas as atenções relacionadas à *escrita* e *análise de dados* devem ser mantidas aqui. Uma vez que a IA pode replicar ou amplificar vieses presentes nos dados de treinamento ou mesmo no desenvolvimento dos modelos, a geração de códigos ineficazes ou até discriminatórios é uma possibilidade. Embora a IA possa escrever código rapidamente, sua qualidade não é garantida, exigindo revisões e ajustes por parte dos programadores. Esta etapa pode resultar, inclusive, em perda de tempo, comprometendo a eficiência da ferramenta (Sohail *et al*., 2024).

Outra questão relevante é a responsabilidade e a prestação de contas. Quando um código gerado pela IA apresenta falhas ou causa problemas, é difícil atribuir claramente a responsabilidade, complicando a identificação de soluções adequadas (Silva, 2020). É preciso um cuidado especial, pois eventualmente a IAG pode gerar códigos que são efetivos em suas tarefas, mas que não são seguros, abrindo brechas de dados de diferentes naturezas e facilitando mesmo ataques virtuais (Liao, Vaughan, 2023; Bail, 2024; Ray, 2023)

Frequentemente, a comunidade de programadores adota uma postura mais aberta e colaborativa, o que implica na disponibilidade de seus códigos em repositórios abertos e gratuitos, como o Github[50] e Gitlab[51], além de contribuírem tirando dúvidas em fóruns como o Stack[52]. Geralmente, essa prática segue a lógica de construir em conjunto sobre as melhores soluções disponíveis. Entretanto, os riscos permanecem os mesmos de outros pontos elencados anteriormente. Há risco do plágio de códigos criados por outras pessoas ou empresas; há riscos de vazamentos e erros de várias naturezas. A questão é tão subestimada, que a Agência de Cibersegurança Francesa (*Agence Nationale de la Sécurité des Systèmes d'Information, ANSSI*) em conjunto com o Escritório Federal de Segurança da Informação da Alemanha (*Bundesamt für Sicherheit in der Informationstechnik, BSI*) lançaram juntos um relatório sobre oportunidades e, especialmente, riscos no uso de LLMs como assistentes de programação, e algumas medidas para mitigação (AI Code Assistants, 2024).

Além dessas questões éticas, há desafios no processo de aprendizado. A dependência excessiva da tecnologia pode inibir o desenvolvimento de habilidades críticas e a resolução criativa de problemas complexos por parte dos pesquisadores (Bail, 2024). A exemplo do estabelecido no tópico da escrita acadêmica, nunca utilize diretamente o código gerado pelas IAG e atente-se às mudanças sugeridas por elas, para que o aprendizado não seja inibido. Prefira usar soluções especializadas na geração e correção de programação[53] e atente-se às regras pelas quais tais sistemas são geridos em termos de privacidade e segurança, já que essas interações também podem ser usadas para treinar tais modelos.

7 – Transcrição

Não encontramos referências explícitas ao uso ético e responsável de tecnologias de IA para a transcrição de áudio e vídeo. O mesmo se aplica para a descrição de imagens visando a acessibilidade ou descrição denotativa da imagem. A transcrição

50. Disponível em: https://github.com/. Acesso em 2 nov. 2024.
51. Disponível em: https://gitlab.com/. Acesso em 2 nov. 2024.
52. Disponível em: https://stackoverflow.com/. Acesso em 2 nov. 2024.
53. Aqui nos referimos a plataformas, aplicativos e soluções, como o Github Copilot, mas também começamos a perceber a criação de modelos especializados em linguagem de programação, a exemplo de Code Llama, Codestral, Star Coder, Cursor.

é certamente uma das tarefas mais trabalhosas e enfadonhas de pesquisas qualitativas e o uso de tais soluções tende a acelerar significativamente o trabalho. Aparentemente, tais ferramentas são menos suscetíveis a alucinações e aos vieses dos modelos de linguagem.

Com base nos testes realizados até o momento, os principais desafios no uso dessas ferramentas derivam da própria qualidade dos áudios e vídeos inseridos nas plataformas. Em outras palavras, se o material carregado apresenta áudio de baixa qualidade, problemas de dicção, sotaques ou dialetos específicos de uma comunidade, entre outros fatores, as ferramentas encontrarão dificuldades em realizar uma transcrição precisa de conversas, entrevistas, entre outros tipos de gravação[54].

⚠ Cuidados

Aqui, *a preocupação maior deve ser com a privacidade dos dados dos participantes da pesquisa.* Ou seja, as diretrizes para assegurar a confidencialidade, anonimato e proteção dos dados dos participantes deve ser a primeira preocupação em seu uso, especialmente se a solução usa os dados fornecidos para o treinamento do próprio modelo. Ao se tratar de dados biométricos, com implicações para a segurança pessoal, **nunca associe arquivos de voz a determinado indivíduo, identificando-o**. Novamente, evite utilizar estas ferramentas para dados sensíveis e inéditos.

A segunda questão importante a ser observada é a possibilidade da geração de imprecisões ou mesmo de erros. Apesar de avançarem rapidamente neste campo, atualmente as transcrições com IA ainda erram com alguma frequência, especialmente, quando há múltiplos falantes (por exemplo, grupo focal ou entrevistas em grupo). Portanto, deve-se validar a transcrição gerada pela máquina em relação ao conteúdo original. Esta tarefa geralmente demandará escutar o áudio gravado na coleta e corrigir eventuais erros apresentados pela ferramenta.

Finalmente, é costumeiro que em pesquisas qualitativas a transcrição inclua outros aspectos para além das falas em si, como o tom de voz e o estado emocional dos participantes. Em outras palavras, eventualmente as pessoas vão usar de ironia ou aumentar o tom para serem mais claros, vão existir momentos de raiva, choro e alegria, entre outros. No atual momento, as IAs não são capazes de detectar este tipo de linguagem não-verbal de maneira confiável; porém, o desenvolvimento futuro dos chatbots aponta para uma comunicação pautada emocionalmente, isto é, para que possam reconhecer e interpretar nuances emocionais, entendendo também o contexto da conversação[55].

8 – Tradução

A tradução automática, baseada em outras tecnologias de Inteligência Artificial, já vinha sendo utilizada no âmbito acadêmico, tanto para o auxílio de leituras, como para a redação. Neste último ponto, destaca seu uso para a elaboração de *abstracts*, isto é, resumos em língua inglesa, exigidos em diversos tipos de produção textual [56].

Contudo, diante do novo cenário das IAs generativas, a maior parte das diretrizes utilizadas para a construção deste guia não menciona ou não proíbe a tradução por Inteligência Artificial Generativa. Existe inclusive um consenso que tal uso por autores que não tem o inglês como língua nativa é uma forma de incentivar a publicação em periódicos de alto impacto e equilibrar o jogo da publicação acadêmica (Cotton *et al*., 2024; Dwivedi *et al*., 2023; Perkins, Roe, 2024; Ray, 2023; Sohail *et al*., 2023).

O treinamento mais abrangente dos grandes modelos de linguagem em inglês atraiu a atenção de pesquisadores em todo o mundo, interessados em traduzir suas línguas nativas para o inglês ou corrigir textos escritos nesse idioma. “Diferentemente do Google Tradutor e similares, as LLMs parecem ser capazes de entregar expressões mais fidedignas, mais similares às nativas e especialmente mais próximas do tom e estrutura de materiais acadêmicos” (Sampaio *et al*., 2024a, p. 12), aparentando uma maior capacidade de detectar as nuances da língua e de soar como um nativo. A

54. Agradecemos a Cristiane Sinimbu por este relato.

55. Esta capacidade foi primeiramente anunciada pela OpenAI através do modelo ChatGPT 4o. Pode-se encontrar uma discussão a respeito em https://medium.com/@fbenbelkhir80/how-does-chatgpt-handle-contextual-understanding-b70911cc677e, embora tenha sido tratada por diversos veículos, em seu devido momento.

56. O uso do Google Tradutor, entendido como principal representante da tradução e uso temático geral, já era discutido num texto de 2015. Além de erros grosseiros, relacionados ao contexto de uso idiomático, a tradução automática geralmente não é capaz de encontrar os termos altamente específicos da linguagem acadêmica. Conferir Sabbatini, Marcelo. “3 dicas para não passar vergonha com o resumo de um artigo científico”. https://www.marcelo.sabbatini.com/3-dicas-para-nao-passar-vergonha-com-o-resumo-de-um-artigo-cientifico/. Acesso 5 nov. 2024.

possibilidade de aprimorar o resultado através de sucessivas iterações também apontam para um texto melhor elaborado.

⚠️ Cuidados

Os cuidados anteriores já estabelecidos na *escrita* devem ser seguidos. **Nunca utilize diretamente a tradução feita por uma IAG sem uma detida análise e correção humana**, sob pena da inclusão ou exclusão de informações presentes no texto original. É importante lembrar que, mesmo em tarefas técnicas como uma tradução, os chatbots podem alucinar e distorcer o texto. São necessários *prompts* adequados para garantir que a tradução não seja literal e que esteja de acordo com a linguagem acadêmica (Lu *et al.*, 2023). Os vícios de linguagem e outros erros já apontados na seção de escrita se repetem aqui.

Atente-se às diretrizes presentes no periódico, associação ou congresso para o qual o material será enviado. Verifique se existe a necessidade de indicar a tradução realizada por Inteligência Artificial. Em caso positivo, siga os princípios gerais de transparência e de autoria para a submissão.

9 – Pareceres e avaliações

Uma classe de atividade que mais demanda esforço intelectual e tempo, porém que não é devidamente reconhecida em termos de prestígio acadêmico, a elaboração de pareceres poderia ser beneficiada pela tecnologia criativa. Por evidente, a elaboração automática ou facilitada por IAG de qualquer tipo de parecer científico (artigos em revisão, trabalhos de conclusão, etc.) também possui um grande apelo para o pesquisador. Atualmente, parece não haver consenso sobre este uso específico. Algumas editoras consideram essa prática uma possível solução para a dificuldade em encontrar bons pareceristas, diante do volume crescente da produção científica (Frontiers, 2020)[57]. Nesta direção, Lopezoza (2023) apresenta 10 propostas ousadas sobre como as IAGs podem ser utilizadas no fluxo editorial desde o recebimento do artigo até sua divulgação, contribuindo para:

1. Verificar se o manuscrito recebido está alinhado com os principais temas da revista.
2. Assegurar que o manuscrito esteja de acordo com os padrões da revista.
3. Verificar se há plágio e identificar qualquer conteúdo potencial gerado por IA.
4. Identificar candidatos qualificados para revisar um manuscrito específico.
5. Elaborar um e-mail para os autores detalhando as propostas de melhoria sugeridas pelos avaliadores.
6. Editar o manuscrito final quanto a erros de gramática e ortografia.
7. Adaptar as citações para que correspondam ao estilo de citação da revista.
8. Projetar materiais visuais, como capas de edições e imagens.
9. Desenvolver uma campanha temática baseada no Twitter para um manuscrito específico.
10. Elaborar comunicados de imprensa abrangentes para um público mais amplo.

Em outro texto, Kankanhalli (2024) faz um resumo dos principais pontos nos quais estes sistemas de inteligência artificial generativa poderiam ajudar e apresenta um quadro, incluindo ferramentas que podem ser úteis em sua avaliação.

57. A utilização de IA para suprir a falta de revisores dispostos a este trabalho, entretanto, pode ser interpretada como uma instância do solucionismo tecnológico. A não-remuneração dos revisores tem sido um dos pontos levantados num debate mais amplo sobre o custo, acesso e distribuição das publicações científicas, com um modelo que beneficia estas mesmas grandes editoras privadas, caracterizadas por lucros extraordinários.

Tabela 1 – Potencial de Automação por IA em Publicações Acadêmicas

Tarefa	Potencial de automação de IA	Exemplo de ferramenta de IA
Verificação de formato: verificação de que o manuscrito segue a regra de formato de estrutura, estilos, referências e metadados da publicação	Alto	Penelope.ai
Detecção de plágio: identificação da extensão e natureza da cópia de outras fontes sem atribuição de fonte	Baixo	iThenticate, ZeroGPT
Qualidade da linguagem: avaliação da legibilidade, coesão e lógica apropriadas para o público	Médio	UNSILO
Correspondência manuscrito-revisor: encontrar revisores adequados para um manuscrito usando o perfil do revisor	Alto	TPMS
Escopo/relevância: avaliação da adequação ao escopo da publicação	Médio	UNSILO, GPT-4
Solidez/rigor: verificação de que a metodologia e análise do estudo são rigorosas e robustas	Médio-Baixo	Enago, StatCheck, StatReviewer
Novidade: novidade ou desvio do corpo existente de conhecimento	Baixo	ReviewAdviser
Significado: importância do fenômeno que o manuscrito está focando	Baixo	ReviewAdviser
Escrita e apresentação: avaliação da clareza, precisão e eficácia da apresentação	Alto	Grammarly, Hemingway Editor
Verificação de reprodutibilidade: codificação do autor e checagem de análise de dados	Alto	GPT-4

Fonte: Kankanhalli (2024, p. 79, tradução nossa, grifos no original)

⚠️ Cuidados

Para além das preocupações mais evidentes com a qualidade, é preciso um cuidado extra com a segurança dos dados provenientes de revistas científicas e seus autores. Como já dito e repetido, a IAG usa o *input* dos usuários para seu treinamento, algo que ocorre a cada interação. Diante de informações sensíveis, como artigos, frutos de pesquisas científicas, que utilizam dados inéditos e originais, estamos diante de um risco de vazamento de dados (União Europeia, 2024).

É importante notar que o uso de ferramentas de IAG no processo de revisão por pares pode levar a uma priorização seletiva de variáveis, métodos e descobertas de certas subdisciplinas e tradições que são dominantes nos dados de treinamento (Susarla *et al*., 2023). Isso pode potencialmente limitar a diversidade e a inovação na pesquisa acadêmica.

Talvez o uso mais seguro no curto prazo seria para o que Kankanhalli (2024) chama de pré-avaliação por pares (*pre-peer review*), que consiste em ações de checagem mais simples, como averiguar o formato, a possível existência de plágio e a qualidade da linguagem. A questão de segurança acima apontada permanece, de qualquer forma.

Assim, deve-se sempre conferir se a proibição ou não do uso de tais ferramentas para a elaboração de pareceres, além de privilegiar modelos abertos ou que não usem as interações para seu treinamento. **Na dúvida, recomenda-se não usar IAG para a redação de pareceres**. *No caso de uso, nunca utilizar diretamente o texto gerado ("copiar e colar") e proceder com sua revisão e edição meticulosa*.

10 – Seleção

O uso da IA Generativa em bancas de seleção, a exemplo de processos seletivos para pós-graduação, seleções e concursos públicos para docentes, ainda é um terreno praticamente inexplorado do ponto de vista da literatura[58]. Contudo, apesar de entendermos que esta é uma aplicação que agrega elementos de outros princípios práticos, sobretudo a escrita e a elaboração de pareceres, sua ocorrência na realidade

58. O mais próximo que encontramos foram artigos tratando sobre memoriais para a aplicação para residências em Medicina, na qual os candidatos precisam escrever sobre suas carreiras e experiências (Chen *et al.*, 2024; Mangold, Ream, 2024) e usaremos o paralelo para alguns casos.

— como já presenciado pelos autores — e suas implicações demandam atenção específica. Como atualmente a maioria dos editais de seleção não contemplam normativas em relação ao uso da IAG, buscaremos suprir esta lacuna por meio de algumas recomendações.

Um primeiro ponto a ser considerado é o uso de IAG para a redação do pré-projeto de pesquisa, memorial, carta de intenções ou qualquer outro elemento escrito pelos candidatos utilizado como etapa de avaliação. Aqui, retomamos as considerações realizadas em relação à escrita acadêmica e à detecção de texto gerado por IAG. Diante da impossibilidade de se afirmar categoricamente a partir das ferramentas de detecção se houve uso de IA, com a possível judicialização do processo, recomendamos uma abordagem indireta. Neste sentido, o produto textual deve ser avaliado em função das limitações do "texto GPT" (Mangold, Ream, 2024), a saber:

- Falta de conexão pessoal e inteligência emocional;
- Falta de coesão e articulação entre ideias;
- Superficialidade na abordagem de questões teórico-metodológicas;
- Ausência de exemplificação;
- Textualização vaga, com excesso de adjetivos e advérbios;
- Referências não verificáveis.

Estes critérios incidem diretamente sobre a avaliação do texto enquanto uma produção original, criativa, e significativa para a seleção em questão. Ao analisarem o uso de IA na criação de memoriais para residência em Medicina, Chen *et al*. (2024) denotam que o uso de apoio externo não surgiu com as IAs Generativas, já que os estudantes já podiam utilizar *softwares* avançados de revisão gramatical, ou até mesmo contratar consultores especializados para revisar ou elaborar os textos.

Ressaltamos que, a exemplo de um texto "copiado e colado", ou produzido por uma terceira pessoa (isto é, "comprado"), fases posteriores do processo seletivo podem ser utilizadas para triangular o domínio do candidato acerca do texto apresentado. Numa entrevista de defesa ou seleção, por exemplo, pode ser solicitado que o candidato discorra sobre as referências apresentadas ou defenda pontos específicos de seu projeto ou memorial. Chen *et al*. (2024), ao falarem de memorial, reforçam que os avaliadores precisam ter plena noção da razão pela qual o material é solicitado. O mesmo vale para processos seletivos. Qual o objetivo de cada material? O que os integrantes de comitês de seleção querem efetivamente avaliar ao pedirem tais manuscritos?

Outro ponto a ser considerado é a utilização de IAG para a avaliação de projetos e outros textos enviados pelo candidato. Para este fim, um avaliador poderia utilizar a capacidade de um agente conversacional para analisar o projeto, inclusive a partir de um *prompt* contendo os critérios estipulados em edital. Ainda que possa ser um uso tentador, no sentido de auxiliar uma tarefa que é trabalhosa e complexa, atentamos para dois princípios básicos discutidos anteriormente. O primeiro diz respeito à propriedade intelectual do candidato, que será compartilhada com o Grande Modelo de Linguagem e com a empresa responsável pela ferramenta. E, segundo, tal uso deve ser acompanhado de agência humana, não substituindo a leitura atenta por parte do avaliador. As implicações de uma avaliação possuem grande repercussão sobre a vida e o futuro do candidato, assim como do papel desempenhado pela universidade; neste sentido, a IA nunca deve ser utilizada em substituição do trabalho de avaliação.

Finalmente, a capacidade dos agentes conversacionais também vislumbra usos como a simulação da defesa de projeto, por parte dos candidatos, antecipando possíveis perguntas a serem realizadas pelos avaliadores. De forma similar, a IA pode auxiliar na leitura e estudo da bibliografia utilizada na confecção de pré-projetos ou que seja utilizada como base para provas escritas. Nestes quesitos não encontramos riscos éticos, sendo que ao enviar seu material para a ferramenta de IA o candidato assume os riscos relacionados à privacidade e direito autoral, além dos limites intrínsecos dessas ferramentas (*i.e.* alucinações, *bias* etc.)

O uso de provas orais tem sido aventado como uma possibilidade para as bancas, notadamente as de processo seletivo, porém notamos que, ao serem realizadas online, isso pode facilitar que os candidatos usem rapidamente as IAG como parceiros nas respostas. Então, é preciso uma criatividade extra por parte dos integrantes do processo seletivo para fazer questionamentos e pedidos aos candidatos que não possam ser facilmente descritos em uma IAG. Logicamente, permanece a competência da banca de ao fim afirmar se o candidato tem domínio ou não dos temas em sua resposta.

Abaixo reproduzimos as sugestões de uso responsável de IAG por candidatos à residência médica, compreendendo que os mesmos princípios se aplicam para a elaboração de projetos de pesquisa e outros materiais solicitados em processos seletivos e concursos públicos acadêmicos.

Quadro 1 - Recomendações de Chen et al. (2024) e Mangold e Ream (2024)

Chen, Bowe, Deng (2024)	Mangold, Ream (2024)
1. Não confie na IAG para criar todo o seu memorial.	1. Siga e respeite as diretrizes específicas de cada programa [ou edital]: antes de enviar os materiais de inscrição, certifique-se de estar ciente das diretrizes de cada programa relacionadas ao uso de IA.
2. Converse com a IAG para debater ideias relevantes para as suas próprias experiências e qualidades pessoais, a fim de ajudá-lo no desenvolvimento do rascunho inicial.	2. Esteja ciente do plágio e mantenha a autenticidade: ao utilizar IA para criar materiais de inscrição, considere que os elementos criados podem ser considerados plágio. Uma inscrição relevante manterá as experiências e a perspectiva autênticas do candidato.
3. Use a IAG para sugerir revisões em seu conteúdo, à medida que você for elaborando o rascunho.	3. Divulgue claramente o uso da IA: em todos os aspectos da inscrição em que for usada, a IA deve ser claramente declarada e divulgada aos avaliadores.
4. Use a IAG para melhorar a legibilidade de escrita, à medida que você se aproxima do documento final (Chen et al, 2024, p. 255).	4. Revise e garanta que os erros sejam corrigidos: é essencial revisar cuidadosamente todos os materiais gerados por IA antes do envio para edição e correção de erros.
	5. Seleção do uso da IA: o candidato deve considerar cuidadosamente quais aspectos da inscrição podem ser passíveis de uso de IA. Evite usar IA em tarefas que devem refletir a percepção e o trabalho originais do candidato (memorial, carta de intenção [projeto de pesquisa]) para maximizar a autenticidade e o impacto desses componentes da inscrição.

Adaptado pelos autores a partir de Chen *et al*. (2024) e Mangold e Ream (2024).

⚠ Cuidados

No caso dos promotores da seleção incluírem no edital regras específicas em relação ao uso da IA, recomendamos que os princípios éticos gerais abordados neste guia sejam utilizados. Contudo, reforçamos a necessidade de que os critérios utilizados para a caracterização desse uso sejam claros e verificáveis. No caso de ferramentas de detecção, não existe na atualidade transparência algorítmica, no sentido de garantir que os resultados sejam confiáveis e diversos testes têm evidenciado a falibilidade da tecnologia. Um único "falso positivo", isto é, a sinalização de uso de IA num texto escrito por um humano, pode comprometer a idoneidade de todo processo seletivo (ver o tópico sobre detecção de IA). Chen *et al.* (2024, p. 255) fazem as seguintes recomendações ao comitê de seleção.

1. Não presuma que um candidato não usou o IAG para elaborar sua candidatura.
2. Reconheça que os métodos atualmente disponíveis para detectar textos gerados por IA são imperfeitos.
3. Examine outros aspectos do aplicativo para corroborar as informações apresentadas nas respostas baseadas em texto.
4. Questione o objetivo do memorial e como ele se relaciona com o processo de seleção de forma holística (Chen et al., 2024, p. 255).

De toda forma, cada processo seletivo terá questões e condições únicas, cabendo a seus integrantes definirem as regras mais adequadas para seu contexto. Em nossa visão, a transparência dessas regras tende a ser a melhor decisão. Como já dito anteriormente, a simples proibição tende simplesmente a incentivar que o uso aconteça em segredo.

11 – Agentes de IA

Especialistas acreditam que as inteligências artificiais generativas tendem a evoluir para agentes de IA[59]. Basicamente, seriam ferramentas (*bots*) treinadas e especializadas para tarefas mais específicas, sendo inclusive capazes de tomar decisões em seu escopo. Dito de outra forma, enquanto o ChatGPT, Claude, Gemini e chatbots similares são pensados como uma ferramenta para diversas tarefas, esses robôs são idealizados como ferramentas para realizar uma ou poucas tarefas. Este debate centra-se, em grande medida, nos agentes de IA, ou agentes inteligentes, que detectam o seu ambiente e tomam medidas para atingir os seus objetivos específicos. Equipados com algoritmos e capacidades de IA, estes agentes podem interpretar dados, tomar decisões e interagir autonomamente com seres humanos ou outros sistemas.

Um exemplo acessível (*no code*) se popularizou quando o ChatGPT 4 permitiu que seus usuários criassem esses agentes (*GPTs*) específicos e os tornassem disponíveis para outras pessoas numa espécie de *GPT Store*. Para além das suas funções individuais, é possível acionar outros agentes na janela de contexto do ChatGPT, então você pode, por exemplo, acionar o GPT que elabora fichamentos, outro especializado em formular análises textuais e depois acionar um outro GPT especializado em criar mapas mentais, tudo sem necessitar de abrir novas conversas. Em caso de livros e arquivos maiores, você teria indicações bem mais aprofundadas de seu conteúdo, que podem ajudar na produção de artigos ou mesmo de aulas.

Em outro exemplo, para fazer uma exploração inicial dos resultados de uma pesquisa baseadas em grupos focais, ficamos com centenas de páginas para análise. O pesquisador poderia usar um GPT especializado em análise qualitativa temática e, em seguida, chamar outro especialista em análise de dados e aplicar análises textuais automatizadas. Um terceiro para realizar visualizações mais pertinentes e daí por diante. Rodrigues *et al.* (2024) fizeram um interessante experimento avaliando vários desses GPTs acadêmicos para a busca e seleção de literatura científica e os resultados iniciais sugerem que o ChatGPT pode ser uma ferramenta complementar útil para os pesquisadores, aumentando a eficiência e permitindo o acesso rápido a informações.

Além desses exemplos, destacamos que João Lima, um pesquisador brasileiro do Direito, criou inúmeros GPTs para explorar a base de Teses e Dissertações da CAPES, permitindo as buscas por diferentes áreas de concentração e de pesquisa, facilitando diversos tipos de pesquisa bibliométricas e mesmo cientométricas[60]. O *Poe*[61] é um outro site dedicado a permitir que os usuários testem diferentes LLMs e criem *bots* para diferentes tarefas, que podem ser usados em conjunto, conforme os comandos do usuário. Como sempre, é preciso cuidado[62].

Atualmente, com o avanço do conhecimento em programação, já se discute a possibilidade de agentes de IA mais autônomos, capazes de tomar essas decisões de forma significativamente independente, terceirizando diversas tarefas da pesquisa e, em certos casos, conduzindo a pesquisa por completo. Como exemplo, este repositório[63] mostra uma ferramenta que utiliza oito agentes de IA, que trabalham juntos para conduzir uma pesquisa em certo tema, do planejamento à publicação. Um outro exemplo que chamou bastante atenção foi a empresa japonesa de IA, denominada *Sakana*, que apresentou um pesquisador autônomo baseado em IA para produzir uma pesquisa[64].

⚠️ Cuidados

Aqui, todos os pontos já apresentados devem ser reforçados em termos de viés, privacidade, propriedade intelectual, plágio, agência humana e afins. Geralmente, as diretrizes não mencionam diretamente agentes de IA e pode-se inferir que as regras para cada *bot* são idênticas aos usos específicos de modelos e outras soluções de IA Generativa. Como usual, *verifique se as regras específicas de privacidade dos agentes não são diferentes das presentes nos modelos de linguagem*.

59. YEE, Lareina; CHUI, Michael; ROBERTS, Roger. Why agents are the next frontier of generative AI. **McKinsey Quartely**, 24 jul. 2024. Disponível em: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/why-agents-are-the-next-frontier-of-generative-ai. Acesso em: 4 out. 2024.

60. LIMA, João. **GPTs Pós*BR - Produções Acadêmicas da Base de Teses e Dissertações da Capes**. Disponível em: https://joaoli13.github.io/GPTs.html. Acesso em: 3 out. 2024.

61. Disponível em: https://poe.com/. Acesso 14 out. 2024.

62. Em sua versão gratuita, por exemplo, as suas interações com os modelos dentro do Poe podem ser usadas para seu treinamento. A política de privacidade do Poe pode ser lida na íntegra em: https://poe.com/privacy. Acesso 14 out. 2024.

63. Disponível em: https://github.com/assafelovic/gpt-researcher/tree/master/multi_agents. Acesso 2 out. 2024.

64. SAKANA AI. The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery. **Sakana.ai**, 13 ago. 2024. Disponível em: https://sakana.ai/ai-scientist/. Acesso em: out. 2024.

12 – Detecção de IA Generativa

"Isto é IA!". Conforme o uso de texto generativo se dissemine e inclusive passe a predominar nos resultados de buscadores da Internet, preocupações com autoria, originalidade e qualidade se estendem. Diante do dilema, algumas empresas afirmam possuir soluções capazes de detectar o uso de Inteligência Artificial Generativa com alta precisão. No entanto, há bastantes relatos indicando que tais tecnologias apresentam falhas significativas (Dalalah, Dalalah, 2023) e podem ser facilmente burladas com poucas edições humanas[65]. Então, **tenha a ciência de que, no atual momento, a maioria dos detectores de IAG, notadamente os gratuitos, não são confiáveis**!

⚠ Cuidados

Apesar de vários detectores indicarem valores sobre a possibilidade de um conteúdo ser gerado por IA, esse número tende a ser apenas um valor aproximado, exibindo grande margem de erro. Por exemplo, detectores podem afirmar que um texto tem 70% de chance de ter sido feito por IA, mas o próprio detector não funciona 100%, então, na prática, os 70% iniciais não são precisos (Spinak 2023a, 2023b). De forma similar, mesmo o uso de soluções técnicas como a aplicação de marcas d'água que imprimem padrões específicos ao texto são suscetíveis à manipulação, levando à falta de confiabilidade dos detectores em cenários práticos (Sadasivan et al., 2023).

Existem duas principais razões para isso. Não sabemos exatamente como esses modelos funcionam internamente, apenas questões gerais de seu funcionamento. Além disso, existe um significativo descompasso entre os níveis de investimento nos modelos de linguagem, frequentemente produzidos pelas *Big Techs,* em comparação com os detectores de IA, geralmente mantidos por empresas menores com foco educacional e acadêmico. Neste cenário, é complexo imaginar como essa conta irá fechar. Ademais, os próprios detectores podem apresentar vieses e outros problemas. Por exemplo, uma pesquisa demonstrou que detectores de IA tendem a apontar incorretamente os textos feitos por autores não-nativos em inglês como sendo gerados por IA (Liang *et al*., 2023), pela própria maneira que estes detectores são treinados nesta língua. Nesse âmbito, os detectores não conseguem distinguir textos criados por IAG de textos apenas revisados com tecnologia similar. Da mesma forma, podemos ter casos indivíduos ou grupos de pessoas que escrevem adotando padrões similares aos Modelos de Linguagem, a exemplo de neurodivergentes, o que pode gerar falsos positivos (Dwivedi *et al*., 2023, p. 27).

Diante dessas limitações tecnológicas, é importante fugirmos da lógica do "dataísmo", do "tecnosolucionismo" (Morozov, 2018; Han, 2022), nas quais sempre precisamos de mais dados, tecnologias maiores e mais avançadas para resolver um problema gerado por estas mesmas tecnologias. Apesar de ter uma série de questões a serem debatidas, por ora *as instituições de pesquisa e ensino devem amparar técnica e legalmente os professores e pesquisadores para fazer essa detecçã*o.

Conforme o estudo de Moorhouse *et al*. (2023), muitas universidades têm adotado *softwares* de detecção como Turnitin e GPTZero, porém o corpo acadêmico tem destacado a necessidade de cautela. A precisão dessas ferramentas é incerta, e elas podem ter problemas de privacidade, sendo sugeridas como ferramentas de apoio, mas não como única evidência de plágio.

Neste sentido, Cotton *et al*. (2024, p. 3-5) sugerem uma abordagem combinada de técnicas automatizadas e manuais de avaliação, incluindo a educação dos estudantes sobre plágio, a exigência de declarações de autoria, a submissão de rascunhos para revisão e a análise cuidadosa de padrões linguísticos, de fontes, originalidade e de precisão factual. Eles também notam que a escrita humana tende a ser mais contextualmente consciente e responsiva às necessidades do público, enquanto o texto gerado por IA tende a ser mais genérico e menos adaptado a contextos específicos. Os textos gerados por IA, embora gramaticalmente corretos, muitas vezes carecem de "estilo", "alma", apresentando-se vagos e planos (Cotton *et al*. 2023; Chen *et al.*, 2024; Mangold, Ream, 2024; Perkins *et al*., 2024). Novamente, todas as questões apontadas no tópico de escrita são válidas para a detecção.

Essas características podem ajudar educadores e pesquisadores na identificação de conteúdo potencialmente gerado por IA. Porém esses

65. Recomendamos uma interessante matéria que discute a temática. COOK, Jodie. AI content detectors don't work (the biggest mistakes they have made). **Forbes**, 4 jul. 2024. Disponível em: https://www.forbes.com/sites/jodiecook/2024/07/04/ai-content-detectors-dont-work-the-biggest-mistakes-they-have-made/. Acesso em: 4 out. 2024.

profissionais precisam ser amparados por diretrizes institucionais para fazerem essas avaliações, além de comitês específicos capazes de julgar estes casos. Para estudantes e jovens pesquisadores, o objetivo não deve ser o de punição, mas o de elucidação dos vários problemas trazidos por estas atitudes. Em casos mais sérios de efetivo plágio e de falsificação, tais comitês devem ser munidos de efetivas capacidades de sanção para evitar o alastramento dessas atitudes antiéticas (Franco *et al*., 2023; Almeida, Nas, 2024).

(Algumas) Considerações Finais

Na atualidade, boa parte das políticas universitárias de IA se concentram exclusivamente na questão da integridade de pesquisa, geralmente destacando os possíveis efeitos negativos da tecnologia, especialmente plágio e fraude (Cotton *et al*., 2023; Velez, Rister, 2024). Embora algumas diretrizes afirmem que as tecnologias possam ser úteis e que os estudantes possam aprender com elas, geralmente não proporcionam exemplos ou explicações de sua utilização em contexto prático. Além disso, tendem a não abordar adequadamente a possibilidade de uma integração da produção entre humanos e máquinas para a geração de produtos acadêmicos diversos, notadamente os textos (Gonçalves, 2023; Velez, Rister, 2024).

Este guia foi, em alguma medida, influenciado por esses regramentos presentes em diferentes políticas e diretrizes ligadas ao ensino e pesquisa, com um forte foco na integridade acadêmica e nas recomendações associadas. Ainda assim, acreditamos que fomos um passo além, ao entrar na prática do uso de tais ferramentas. Conseguimos não apenas dar as recomendações normativas, mas efetivamente apontar ferramentas, usos e especialmente os cuidados necessários para tal.

Entre outras coisas, isso significa que o teor geral deste guia certamente passará a impressão de uma visão negativa sobre a IA, apresentando mais restrições que usos. Como devidamente questionados por colegas[66], a IA, de forma geral, visa poupar tempo na pesquisa, porém se tomarmos todas as precauções sugeridas, ela não fica inexequível? Dito de outra forma, ela só nos ajuda a poupar tempo se aceitarmos tacitamente todos os seus vieses e problemas cegamente? Em grande medida, aceitamos a crítica, pois optamos por sugerir bastante cautela e reflexão em cada uso. Evidentemente, o objetivo aqui é justamente apresentar sugestões e diretrizes. Em especial, o esforço está em apresentar problemas, questões e limites de tais ferramentas, como forma de alerta a pesquisadores que não estudam ou atuam na área, buscando assim evitar um uso apenas técnico e pouco reflexivo. Na prática, os diferentes pesquisadores deverão fazer as adequações para seus contextos e objetivos específicos.

Por outro lado, assumimos uma visão realista, na qual o uso de tais tecnologias irá acontecer, gostemos ou não. Assim, defendemos que só podemos efetivamente analisar e criticar tais tecnologias se nos apropriarmos adequadamente delas! Em alguma medida, isso quer dizer que praticamente não tratamos sobre a recusa e o não uso da IA, correndo o risco de se produzir um discurso que legitime a sua adoção a todo custo, concedendo validade e inevitabilidade ao uso da tecnologia. Embora essa crítica seja válida, ela não está no escopo deste trabalho, que se concentra na apresentação de diretrizes para o uso ético e responsável. Portanto, trata-se de reflexões vitais a serem feitas pela academia em espaços mais adequados[67].

Atualmente, como dito na introdução, parece que vivemos um sentimento de "policiamento", no qual o uso de IA é visto apenas como abuso ou como uma indicação de limitação intelectual (Velez, Rister, 2024). Os debates parecem acontecer apenas em grandes termos conceituais e normativos, com pouco espaço para uma discussão aplicada. Como sabemos, os modelos mentais dos usuários afetam diretamente como avaliamos os Modelos de Linguagem; se não os conhecemos adequadamente, podemos ter usos equivocados, perigosos, desconsiderando a privacidade, questões de segurança e atuando com confiança excessiva (Liao, Vaughan, 2023). Logo, é importante que docentes e pesquisadores se apropriem das ferramentas e compreendam como elas funcionam.

Da mesma forma, pela própria natureza das IAG, um debate colaborativo parece ser o melhor caminho para compreender dificuldades, frustrações, receios, assim como potencialidades, benefícios e usos pertinentes e inovadores. Um debate que deve incluir toda a comunidade acadêmica, técnicos administrativos, docentes e discentes (Velez, Rister, 2024).

Este guia surgiu de uma inquietação dos pesquisadores pela falta de mais discussões e, especialmente, pela ausência de diretrizes institucionais. Longe de pretender ser definitivo, esperamos que seja um pontapé inicial. Para

66. Agradecemos aos colegas do grupo de pesquisa Margem da UFMG por esta e outras reflexões específicas.

67. Ver o exemplo de Lindebaum e Fleming (2024), que simplesmente sugerem que o ChatGPT não seja usado na pesquisa acadêmica, pois avaliam que os ganhos não superam as perdas.

compreender nossas especificidades, incluindo o cenário e as necessidades de nossos pesquisadores, é preciso parar de ignorar o problema da Inteligência Artificial Generativa ou simplesmente vilipendiá-la. Precisamos começar a realizar pesquisas e discussões mais profícuas, parar de ignorar o "elefante na sala" e trazer a questão para o centro do debate.

As universidades são *loci* naturais de tais discussões e devem incentivar eventos, assim como formar comitês de ética específicos para a questão da Inteligência Artificial Generativa. Discentes e jovens pesquisadores, em especial, precisam encontrar espaços abertos e seguros para exporem suas dúvidas e mesmo apresentar seus objetivos e os bons usos das ferramentas. Devemos pensar em debates ampliados nos quais docentes e pesquisadores sêniores podem enfatizar a importância da autonomia no processo da pesquisa. Precisamos discutir as habilidades necessárias para trabalhar de forma íntegra e ética, assim como devemos estar abertos para soluções e usos criativos da IAG. Institucionalmente, precisamos estar preparados para "fornecer orientações sobre estratégias para identificar e avaliar os riscos de ferramentas habilitadas por IA e estabelecer mecanismos de governança apropriados para a instituição se adequar às políticas e normas para uso de IA" (Almeida, Nas, 2024, p. 24).

Se as instituições nacionais de fomento à pesquisa estão se esquivando da questão, há muitas instâncias nas quais isso pode acontecer. O exemplo de associações científicas, como a Intercom nos permitiu fazer aqui, é uma direção. Da mesma maneira, cada área de pesquisa, por meio de seus fóruns, redes de editores científicos e de pesquisas, entre outras instâncias, pode organizar comitês e grupos de trabalho para elaborar suas próprias diretrizes. Rascunhos podem ser publicizados para receberem colaborações coletivas de seus integrantes, fomentando o debate pelos próprios pesquisadores, sempre respeitando as necessidades e especificidades de cada área. Sintetizando,

> Nesse primeiro nível, as organizações de pesquisa têm oportunidade de estabelecer regras e normas, quando ainda não se tem no país legislações aprovadas por Congresso e governo, que usualmente necessitam de mais tempo para formar arranjos políticos requeridos para aprovação de legislações. Esse deve ser o ponto de partida das instituições de pesquisa e ensino superior no Brasil para avançar responsavelmente no uso e aplicação das tecnologias de IA (Almeida, Nas, 2024, p. 26).

Os princípios e guias são importantes, mas só serão efetivos se combinados "com práticas de governança que ajudem a guiar o uso de ferramentas de IA em projetos de pesquisa científica, desde o design até a produção e publicação de resultados de pesquisa" (Almeida, Nas, 2024, p. 24). Assim, o papel da governança é articular as repercussões da IA sobre indivíduos, comunidades e instituições.

E, claro, tudo isso será melhor amparado se tivermos boas pesquisas acontecendo. Ao longo deste manual, ficou claro que há muito sendo produzido pelos "suspeitos usuais": os grandes centros de pesquisa no mundo, principalmente dos Estados Unidos e da Europa. Ou seja, precisamos de mais pesquisas. Aqui, estamos nos referindo a pesquisas quantitativas de opinião (Pereira *et al*., 2024), grupos focais (Casimiro, 2023), estudos de casos (Saraiva Jr., 2024), relatos de experiência (Divino, 2024; Peres, 2024), revisões bibliográficas (Grossi *et al*., 2024; Jesus, Santarém Segundo, 2024; Rodrigues *et al*., 2024). Mesmo reflexões (Alves, 2023; Ramos, 2023; Barreto, Ávila, 2024; Fonseca, Campiglia, 2024; *Grossi et al.*, 2024; Limongi, 2024; Sampaio *et al*., 2024a) devem ser produzidas com maior frequência, para não reproduzirmos apenas as regras advindas do Norte Global e para que possamos estabelecer diretrizes para o uso responsável, ético e soberano da Inteligência Artificial Generativa no Brasil.

Palavras finais: além do *hype*, em busca do uso soberano de IA Generativa pelo Brasil

Desde o lançamento do ChatGPT, grande parte das maiores discussões sobre os impactos das tecnologias sobre a sociedade tem se focado na Inteligência Artificial[68]. Não obstante, enquanto acadêmicos, precisamos adotar uma perspectiva crítica e não ser suscetíveis ao "*hype*" trazido por tais tecnologias[69]. Apresentamos três mitos[70] que acreditamos que não podem ser reproduzidos sem maior reflexão.

O principal alarde a ser evitado é a conexão espúria entre a automatização de tarefas permitidas por essas soluções e a lógica de maior tempo e foco para atividades mais importantes e criativas. De fato, há indicativos de efeitos instrumentais das IAs na pesquisa, que podem ajudar a gerar resultados mais substantivos em determinadas áreas (Silva *et al*., 2024), entretanto, na prática, é mais provável que teremos a precarização de diversas profissões, o que certamente irá afetar o mundo acadêmico. *Não teremos mais tempo na prática, mas precisaremos realizar mais tarefas no mesmo prazo*.

Historicamente, houve uma separação mais clara entre o desenvolvimento tecnológico no Vale do Silício e a pesquisa acadêmica. Agora, essa fronteira está se dissolvendo, pois justamente as maiores empresas de tecnologia estão na vanguarda do desenvolvimento de tais ferramentas, com aplicações diretas e imediatas na pesquisa científica (Liao, Vaughan, 2023). O modelo tradicional de desenvolvimento de *software* acadêmico, por sua vez, era caracterizado por longos ciclos de atualização, realizado por empresas consideravelmente menores. Em contraste, o cenário atual da IA apresenta atualizações significativas a cada poucos meses, às vezes até semanas. Na prática, isso tende a acelerar ainda mais o tempo da pesquisa.

O segundo mito imposto pelo Vale do Silício é de um "tecnosolucionismo", pelo qual as soluções sempre demandam por mais dados e modelos maiores e mais avançados de tecnologia (Morozov, 2018; Han, 2022). Em outras palavras, cada nova tecnologia resolve ou facilita determinadas tarefas e o enfrentamento de alguns obstáculos, mas simultaneamente elas causam novos problemas. Este discurso vai defender que apenas mais tecnologia poderá resolver os novos problemas causados, sendo que na prática frequentemente as melhores soluções passariam por resoluções humanas ou simplesmente por não utilizar essas ferramentas. Muitas vezes poderíamos simplesmente concluir que os ganhos trazidos por certas tecnologias não compensam os problemas trazidos. Por exemplo, já sabemos que os centros de dados necessários para treinar e manter tais modelos geram grandes custos ambientais (Berthelot *et al.*, 2024). Geralmente, se fala de IA para o bem e de benefícios gerais de IA, mas frequentemente se confundem ganhos trazidos com todas as técnicas de IA preditiva com as inteligências artificiais generativas.

O terceiro mito a ser combatido é que essas tecnologias poderão ter um efeito positivo para países do sul global, como o Brasil. Para além do que já sabemos sobre colonialismo de dados (Cassino, Souza, Silveira, 2021; Oliveira, Neves, 2023), Filgueiras e Gomide (2024) sugerem o uso do conceito de "capitalismo de monopólio intelectual" de Cecília Rikap, no qual

> a apropriação de dados e inovações por poucas corporações estrangeiras perpetua a dependência tecnológica e econômica de países em desenvolvimento, como o Brasil. Essa concentração da infraestrutura, na qual a IA existe e opera com dados e algoritmos, faz com que estas corporações se apropriem do conhecimento gerado, transformando, por

68. Um interessante estudo foi conduzido pela FGV Rio com base em 365 mil posts de redes sociais no início de 2024, capturando como a IA é debatida no contexto brasileiro. A pesquisa destaca uma percepção utilitarista e otimista da IA, especialmente em setores de serviços e na administração pública. FGV COMUNICAÇÃO RIO. Imaginários sobre a Inteligência Artificial no debate brasileiro. **FGV Comunicação Rio**, 1 jul. 2024. Disponível em: https://midiademocracia.fgv.br/estudos/imaginarios-sobre-inteligencia-artificial-no-debate-digital-brasileiro. Acesso em: 2 nov. 2024.

69. Uma boa indicação é o podcast "Tech won't save us" do ativista Paris Marx, no qual ele debate várias temáticas buscando desmistificar diversos mitos promovidos por empresas e agentes do Vale do Silício. Disponível em: https://open.spotify.com/episode/4Goe1Q1x8B9vZiFRmhu6O9. Acesso 15 out. 2024.

70. Dora Kaufman (2022) reuniu anos de colunas a respeito de inteligência artificial e seus impactos na pesquisa e na sociedade como um todo. O sugestivo título de "desmistificando a inteligência artificial" nos sugere que há muitos mitos a serem desconstruídos.

> sua vez, as dinâmicas de poder econômico e político global em detrimento dos países em desenvolvimento (Filgueiras, Gomide, 2024, sp.).

Rikap (2021) argumenta que os monopólios intelectuais do século XXI se sustentam pela concentração permanente e crescente de ativos intangíveis, como conhecimento e dados, criando um ciclo de acumulação que perpetua desigualdades globais. Esses monopólios não necessariamente dominam os mercados em que operam, mas monopolizam o conhecimento, transformando-o em ativos de valor econômico exclusivo. Essa capacidade de depredar o conhecimento produzido por outras organizações permite que grandes corporações, como as do setor de tecnologia e farmacêutico, mantenham sua posição dominante e extraiam rendas intelectuais de suas redes de inovação.

A predação e o rentismo são os mecanismos que fundamentam essa concentração de poder. As corporações intelectuais planejam e organizam a produção e a inovação, subordinando outras empresas, universidades e organizações públicas de pesquisa, que passam a depender do acesso ao capital e aos mercados controlados por esses monopólios (Rikap, 2021).

Em outras palavras, o maior perigo do uso dessas ferramentas não está no viés ou alucinações das IAG, mas sim em entregar o conhecimento de ponta produzido no país[71]. Sem as diversas cautelas aqui explicitadas, estaremos constantemente treinando e aperfeiçoando estes modelos que pertencem a grandes empresas que não têm qualquer interesse em compartilhar tais tecnologias. Portanto, é necessário um cuidado redobrado nas parcerias estabelecidas com essas empresas, que, embora possam realizar investimentos localmente, estes são mínimos em comparação ao montante investido em seus países de origem e ainda menores em relação aos seus lucros.

Além de promover o uso ético e responsável da Inteligência Artificial, é fundamental que o Brasil valorize sua soberania nacional e tecnológica nesse processo. Duas importantes associações científicas, a Associação Brasileira de Ciências (ABC, 2024) e a Sociedade Brasileira de Computação (SBC, 2024), apresentaram recomendações para avançar o desenvolvimento da IA no país, destacando a necessidade de investimento em recursos humanos, como programas de bolsas para estudantes que fomentem a inovação nas universidades e colaboração entre academia e setor privado. Ademais, as recomendações enfatizam o potencial da IA em áreas críticas como educação, saúde e meio ambiente, promovendo competitividade e inovação por meio de novos produtos e serviços. Também é destacada a importância de uma política de acesso a bases de dados, com a criação de uma rede nacional de centros de dados coordenada por uma instituição federal, assegurando a qualidade e consistência dos dados utilizados em IA.

O país parece ter "acordado" na proposição do Plano Brasileiro de Inteligência Artificial (PBIA) 2024-2028[72], que destaca importantes investimentos em infraestrutura, capacitação e desenvolvimento tecnológico de IA no país, entretanto é necessário um esforço coletivo da ciência brasileira para desafiar os paradigmas tecnológicos e mesmo científicos já consolidados.

Em outras palavras, a criação de modelos abertos de IAG, soberanos e desenhados para a nossa realidade e para a nossa necessidade deve permanecer como meta, notadamente no campo da pesquisa científica de ponta. A universidade de Stanford já criou o modelo Storm para fins exclusivamente acadêmicos, como exemplo. Não obstante, já há exemplos no Brasil, como a Maritaca AI[73], que foi gestada na Unicamp, que evidenciam que há sim caminhos a serem seguidos pela ciência brasileira.

---

71. Agradecemos especialmente a Daniel Fugisawa por este alerta.

72. BRASIL. Ministério da Ciência, Tecnologia e Inovações. Plano Brasileiro de Inteligência Artificial (PBIA) 2024. **Laboratório Nacional de Computação Científica - LNCC**, 7 ago. 2024. Disponível em: https://www.gov.br/lncc/pt-br/assuntos/noticias/ultimas-noticias-1/plano-brasileiro-de-inteligencia-artificial-pbia-2024-2028 . Acesso em: 3 out. 2024.

73. Macedo, Leandro. MariTalk: Pesquisadores da Unicamp Criam "ChatGPT" e "DALL-E" Brasileiros. https://multitec.com.br/maritalk/. Acesso em 4 nov. 2024.

Declaração do uso de inteligência artificial generativa

Este manual foi inteiramente elaborado pelos autores, mas contou com o auxílio de inúmeras ferramentas de inteligência artificial generativa. SciSpace, Perplexity e Elicit (versões não claras, acesso em setembro e outubro de 2024) foram usados como ferramentas de busca semântica baseada em inteligência artificial para encontrar materiais específicos de organizações, universidades e afins. ChatGPT e Claude foram usados em vários momentos para a exploração de ideias. ChatGPT (versões 4o, 4o com Canvas), Claude (versão 3.5 Sonnet) e Text Cortex (sem versão) foram usados entre setembro e novembro de 2024 para fazer resumos de anotações humanas e para a revisão textual. Após o uso destas ferramentas, os autores revisaram e editaram o conteúdo em conformidade com o método científico e assumem total responsabilidade pelo conteúdo da publicação.

Referências

ACADEMIA BRASILEIRA DE CIÊNCIAS (ABC). **Recomendações para o avanço da inteligência artificial no Brasil.** GT-IA/coordenador: Virgílio Augusto Fernandes Almeida. Rio de Janeiro: Academia Brasileira de Ciências, 2023. Disponível em: https://www.abc.org.br/wp-content/uploads/2023/11/recomendacoes-para-o-avanco-da-inteligencia-artificial-no-brasil-abc-novembro-2023-GT-IA.pdf. Acesso em: 2 out. 2024.

AI CODE ASSISTANTS. Relatório conjunto entre Federal Office for Information Security (BSI) e Agence nationale de la sécurité des systèmes d'information. Bonn, Bundesamt für Sicherheit in der Informationstechnik, 2024. Disponível em: https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/KI/ANSSI_BSI_AI_Coding_Assistants.pdf. Acesso em: 2 out. 2024.

ALEXANDER, A.et al. **Ghosts, robots, automatic writing**: an AI level study guide. Cambridge: Cambridge University, 2021. Disponível em: https://www.cdh.cam.ac.uk/wp-content/uploads/2022/05/ghostfictions_booklet.pdf

ALKAISSI, Hussam; McFARLAN, Samy. Artificial hallucinations in ChatGPT: implications in scientific writing. **Cureus**, San Francisco, CA, v. 15, n. 2, e35179, 2023. DOI: https://doi.org/10.7759/cureus.35179.

ALMEIDA, Virgílio; MENDONÇA, Ricardo Fabrino; Filgueiras, Fernando. ChatGPT: tecnologia, limitações e impactos. **Ciência Hoje**, Rio de Janeiro, [CH396], mar. 2023. Disponível em: https://cienciahoje.org.br/artigo/chatgpt-tecnologia-limitacoes-e-impactos/. Acesso 2 nov. 2024.

ALMEIDA, Virgílio; NAS, Elen. Desafios da IA responsável na pesquisa científica. **Revista USP**, São Paulo, n. 141, p. 17–28, 5 jun. 2024. DOI: https://doi.org/10.11606/issn.2316-9036.i141p17-28.

ALVERO, A. J. *et al*. Large language models, social demography, and hegemony: comparing authorship in human and synthetic text. **Journal of Big Data**, v. 11, n. 1, p. 1-28, 2024. Disponível em: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-024-00986-7.

ALVES, Lynn (org.) **Inteligência Artificial e Educação:** refletindo sobre os desafios contemporâneos. Salvador: Edufba, 2023. Disponível em: https://repositorio.ufba.br/bitstream/ri/38646/1/Inteligência%20artificial%20e%20educação-repositorio.pdf.

ANPD - AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Guia orientativo:** Tratamento de dados pessoais para fins acadêmicos e para a realização de estudos e pesquisas. Brasília, jun./2023. Disponível em: https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/. Acesso em: 31 out. 2024.

ASSOCIATION OF RESEARCH LIBRARIES / COALITION FOR NETWORKE INFORMATION. ARL/CNI AI Scenarios: AI-Influenced Futures, Washington, DC QWest Chester, PA: Association of Research Libraries, Coalition for Networked Information, 2024. Disponível em; https://doi.org/10.29242/report.aiscenarios2024

BAIL, Christopher. Can Generative AI improve Social Science? *In:* **Proceedings of the National Academy of Sciences**, v. 121, n. 21, p. e2314021121, 2024. DOI: https://doi.org/10.1073/pnas.2314021121.

BARRETO, Alana Maria Passos; ÁVILA, Flávia. A Inteligência Artificial diante da integridade científica: um estudo sobre o uso indevido do ChatGPT. **Revista Direitos Culturais**, v. 18, n. 45, p. 91-106, 2023. DOI: https://doi.org/10.31512/rdc.v18i45.1373.

BERTHELOT, Adrien et al. Estimating the environmental impact of Generative-AI services using an LCA-based methodology. **Procedia CIRP,** v. 122, p. 707-712, 2024. DOI: https://doi.org/10.1016/j.procir.2024.01.098

BLOCK, Joern; KUCKERTZ, Andreas. What is the future of human-generated systematic literature reviews in an age of artificial intelligence?. **Management Review Quarterly**, p. 1-6, 2024. DOI: https://doi.org/10.1007/s11301-024-00471-8.

BOYD-GRABER, Jordan; OKAZAKI, Naoaki; ROGERS, Anna. **ACL 2023 policy on AI writing assistance**. ACL 2023, 10 jan. 2023. Disponível em: https://2023.aclweb.org/blog/ACL-2023-policy. Acesso em: 2 dez. 2023.

BSHARAT, Sondos Mahmoud; MYRZAKHAN, Aidar; SHEN, Zhiqiang. Principled instructions are all you need for questioning llama-1/2, gpt-3.5/4. **arXiv preprint** arXiv:2312.16171, 2023. DOI: https://doi.org/10.48550/arXiv.2312.16171.

BUBECK, Sébastien *et al*. Sparks of Artificial General Intelligence: early experiments with GPT-4. **arXiv preprint***, 2023.* DOI: https://doi.org/10.48550/ARXIV.2303.12712.

CAMBRIDGE. **Research publishing Ethics Guidelines for**

journals. 2023. Disponível em: https://www.cambridge.org/core/services/authors/publishing-ethics/research-publishing-ethics-guidelines-for-journals/authorship-and-contributorship#ai-contributions-to-research-content. Acesso em: 2 dez. 2023.

CASIMIRO, Adelaide Helena Targino. **Tecnologias pós-humanistas e o mercado de trabalho brasileiro para Arquivistas***:* percepções e desafios por meio de estudos de cenários prospectivos. Tese (Doutorado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Universidade Federal da Paraíba (UFPB), João Pessoa, 2023. Disponível em: https://repositorio.ufpb.br/jspui/handle/123456789/31342.

CASSINO, João Francisco; SOUZA, Joyce; SILVEIRA, Sérgio Amadeu da (org.). **Colonialismo de dados:** como opera a trincheira algorítmica na guerra neoliberal. São Paulo: Autonomia Literária, 2022.

CHEN, Jenny X.; BOWE, Sarah; DENG, Francis. Residency Applications in the Era of Generative Artificial Intelligence. **Journal of Graduate Medical Education**, v. 16, n. 3, p. 254-256, 2024. DOI: https://doi.org/10.4300/jgme-d-23-00629.1.

CHETWYND, Ellen. Ethical use of artificial intelligence for scientific writing: Current trends. **Journal of Human Lactation**, v. 40, n. 2, p. 211-215, 2024. DOI: https://doi.org/10.1177/08903344241235160.

CONG-LEM, Ngo; SOYOOF, Ali; TSERING, Diki. A systematic review of the limitations and associated opportunities of ChatGPT. **International Journal of Human–Computer Interaction**, p. 1-16, 2024. DOI: https://doi.org/10.1080/10447318.2024.2344142.

COPE. **Authorship and AI tools**. COPE, 13 fev. 2023. Disponível em: https://publicationethics.org/cope-position-statements/ai-author. Acesso em: 2 dez. 2023.

COPE, Bill; KALANTZIS. Mary. Generative AI comes to school (GPT and all that fuss): what now?. **Educational Philosophy and Theory**, n. 13-17, 2023. Disponível em: https://doi.org/10.1080/00131857.2023.2213437.

COTTON, Debby R. E.; COTTON, Peter. A.; SHIPWAY, J. Reuben. Chatting and cheating: ensuring academic integrity in the era of ChatGPT. **Innovations in Education and Teaching International,** v. 61, n. 2, p. 228-239, 2023. DOI: https://doi.org/10.1080/14703297.2023.2190148.

CRUZ, Robson Nascimento da. **Bloqueio da escrita acadêmica:** caminhos para escrever com conforto e sentido. Belo Horizonte: Artesã, 2020.

CURTIS, Nigel. To ChatGPT or not to ChatGPT? The impact of Artificial Intelligence on academic publishing. **The Pediatric Infectious Disease Journal**, v. 42, n. 4, p. 275, 2023. DOI: https://doi.org/10.1097/inf.0000000000003852.

DABIS, Attila; CSÁKI, Csaba. AI and Ethics: investigating the first policy responses of Higher Education institutions to the challenge of Generative AI. **Humanities and Social Sciences Communications**, v. 11, n. 1, p. 1-13, 2024. DOI: https://doi.org/10.1057/s41599-024-03526-z.

DALALAH, Doraid; DALALAH, Osama M. A. The false positives and false negatives of Generative AI detection tools in education and academic research: the case of ChatGPT. **The International Journal of Management Education**, v. 21, n. 2, p. 100822, 2023. DOI: https://doi.org/10.1016/j.ijme.2023.100822.

DIVINO, Sthéfano. Inteligência Artificial Generativa no Ensino Superior: diretrizes para superação dos dilemas didáticos, éticos e legais. **Revista Pedagogía Universitaria y Didáctica Del Derecho**, v. 11, n. 1, p. 6-30, 2024. DOI: https://doi.org/10.5354/0719-5885.2024.74070.

SOUZA, Estefane Domingos de; SABBATINI, Marcelo. Integridade Acadêmica em Xeque: Dilemas Éticos da Inteligência Artificial Generativa na Educação e Pesquisa. *In*: I CIFOP - Congresso Internacional Sobre a Formação do Pesquisador [no prelo]. **Anais**... Curitiba: CIDES/PUC-PR, 2024.

DWIVEDI, Yogesh K. *et al*. "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. **International Journal of Information Management**, v. 71, n. 0268-4012), p. 102642, 2023. DOI: https://doi.org/10.1016/j.ijinfomgt.2023.102642.

ELSEVIER. **The use of generative AI and AI-assisted technologies in the review process for Elsevier**, 2023a. Disponível em: https://www.elsevier.com/about/policies-and-standards/the-use-of-generative-ai-and-ai-assisted-technologies-in-the-review-process. Acesso em: 12 dez. 2023.

ELSEVIER. **The use of generative AI and AI-assisted technologies in writing for Elsevier,** 2023b. Disponível em: https://www.elsevier.com/about/policies-and-standards/the-use-of-generative-ai-and-ai-assisted-technologies-in-writing-for-elsevier. Acesso em: 12 dez. 2023.

FCHSSALLA - FÓRUM DE CIÊNCIAS HUMANAS, SOCIAIS, SOCIAIS APLICADAS, LINGUÍSTICA, LETRAS E ARTES. **Diretrizes para a ética na pesquisa e a integridade**

científica: Grupo de Trabalho de Ética em Pesquisa (2022-2023). Frederico Garcia Fernandes (coord.). Brasília: Centro de Gestão e Estudos Estratégicos, 2024. Disponível em: https://anpocs.org.br/wp-content/uploads/2024/03/2024-03-DIRETRIZES-DE-ETICA-NA-PESQUISA.pdf.

FRONTIERS. Artificial Intelligence to help meet global demand for high-quality, objective peer-review in publishing. **Frontiers Science News**, 1 jul. 2020. Disponível em: https://blog.frontiersin.org/2020/07/01/artificial-intelligence-to-help-meet-global-demand-for-high-quality-objective-peer-review-in-publishing/. Acesso em: 12 dez. 2023.

FONSECA, Tina; CAMPIGLIA, Lucila. APREND.AI: Inteligência Artificial Generativa no Ensino Superior. **TECCOGS: Revista Digital de Tecnologias Cognitivas**, n. 28, p. 87–107, 2023. DOI: http://dx.doi.org/10.23925/1984-3585.2023i28p87-107.

FRANCO, Diego; VIEGAS, Luís Eduardo; RÖHE, Anderson. Guia Ético para a Inteligência Artificial Generativa no Ensino Superior. **TECCOGS: Revista Digital de Tecnologias Cognitivas,** n. 28, p. 108–117, 2023. DOI: http://dx.doi.org/10.23925/1984-3585.2023i28p108-117.

FROMMHERZ, Y.; LANGENHORST, J. Digital skills for Humanities and Social Science students: benefits of a blended learning format for teaching programming skills. **Lessons Learned**, v. 2, n. 1, 2022. Disponível em: https://ll.journals.qucosa.de/ll/article/download/37/86.

GAO, Zhengjie *et al*. A Brief Survey on safety of Large Language Models. **Journal of Computing and Information Technology**, v. 32, n. 1, p. 47-64, 2024. DOI: http://dx.doi.org/10.20532/cit.2024.1005778.

GATRELL, C., Muzio, D., Post, C., & Wickert, C. Here, there and everywhere: On the responsible use of artificial intelligence (AI) in management research and the peer-review process. Journal of Management Studies, v. 61, n. 3, p. 739-751, 2024. DOI: https://doi.org/10.1111/joms.13045

GIRARDI, Luana da Silva; PASE, André Fagundes. F. Em busca do comando ideal: um duplo olhar sobre a Inteligência Artificial Generativa na Comunicação Organizacional. **Organicom**, v. 21, n. 44, p. 71-84, 2024. DOI: https://doi.org/10.11606/issn.2238-2593.organicom.2024.220414.

GONÇALVES, Renato. **Criatividade e Inteligência Artificial**. São Paulo: Estação das Letras e Cores, 2023.

GROSSI, Márcia Gorett Ribeiro *et al*. Inteligência Artificial e o modelo ChatGPT: o que as pesquisas estão revelando e um recorte com contexto educacional. **Caderno Pedagógico**, v. 21, n. 7, e5918, 2024. DOI: https://doi.org/10.54033/cadpedv21n7-193.

HABIB, Sabrina *et al*. How does generative Artificial Intelligence impact student creativity? **Journal of Creativity**, v. 34, n. 1, p. 100072, 2024. DOI: https://doi.org/10.1016/j.yjoc.2023.100072.

HACKER, Philipp; ENGEL, Andreas; MAUER, Marco. Regulating ChatGPT and other Large Generative AI Models. 2023 ACM Conference on Fairness, Accountability, and Transparency. *In*: **FACCT '23: THE 2023 ACM Conference on Fairness, Accountability and Transparency.** Chicago: ACM, 2023. DOI: https://dl.acm.org/doi/10.1145/3593013.3594067.

HAMILTON, Leah *et al*. Exploring the use of AI in qualitative analysis: a comparative study of guaranteed income data. **International Journal of Qualitative Methods**, v. 22, p. 1-13, 2023. DOI: https://doi.org/10.1177/16094069231201504.

HAN, Byung-Chul. **Infocracia***:* digitalização e a crise da democracia. São Paulo, Vozes, 2022.

HASSAN, Mahadi; KNIPPER, Alex; SANTU, Shubhra Kanti Karmaker. ChatGPT as your personal data scientist. ***arXiv preprint*** arXiv:2305.13657, 2023. DOI: https://doi.org/10.48550/arXiv.2305.13657.

ICMJE. **Defining the role of authors and contributors**, 2023. Disponível em: https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html. Acesso em 06 nov. 2024.

JESUS, Ananda Fernanda de; SANTARÉM SEGUNDO, José Eduardo. Aplicando o ChatGPT na condução de revisões sistemáticas da literatura: protocolo de pesquisa. **Ciência da Informação**, v. 53, n. Especial, p. 27–38, 2024. DOI: https://doi.org/10.18225/ci.inf.v53i.6666.

KANKANHALLI, Atreyi. Peer review in the Age of Generative AI. **Journal of the Association for Information Systems**, v. 25, n. 1, p. 76-84, 2024. DOI: http://doi.org/10.17705/1jais.00865.

KAUFMAN, Dora. **Desmistificando a Inteligência Artificial**. Belo Horizonte: Autêntica, 2022.

KHAN, Nauman *et al*. Global insights and the impact of generative AI-ChatGPT on multidisciplinary: a systematic review and bibliometric analysis. **Connection Science**, v. 36, n. 1, p. 2353630, 2024. DOI: https://doi.org/10.1080/09540091.2024.2353630.

KOBAK, Dmitry *et al*. Delving into ChatGPT usage in ac-

ademic writing through excess vocabulary. **arXiv preprint** arXiv:2406.07016, 2024. DOI: https://doi.org/10.48550/arXiv.2406.07016.

KORINEK, Anton. Generative AI for economic research: use cases and implications for economists. **Journal of Economic Literature**, v. 61, n. 4, p. 281-1317, 2023. DOI: https://www.doi.org/10.1257/jel.20231736.

LEHR, Steven A. *et al*. ChatGPT as research scientist: probing GPT's capabilities as a research librarian, research ethicist, data generator, and data predictor. *In*: **Proceedings of the National Academy of Sciences**, v. 121, n. 35, p. e2404328121, 2024. DOI: https://doi.org/10.1073/pnas.2404328121.

LEMOS, André. Datificação da vida. **Civitas-Revista de Ciências Sociais**, v. 21, p. 193-202, 2021. DOI: https://doi.org/10.15448/1984-7289.2021.2.39638.

LEMOS, André. Erros, falhas e perturbações digitais em alucinações das IA generativas: Tipologia, premissas e epistemologia da comunicação. **MATRIZes**, v. 18, n. 1, p. 75-91, 2024. DOI: http://dx.doi.org/10.11606/issn.1982-8160.v18i1p75-91.

LEVENE, Alysa. **Artificial intelligence and authorship**. COPE, 23 fev. 2023. Disponível em: https://publicationethics.org/news/artificial-intelligence-and-authorship. Acesso em: 2 dez. 2023.

LIANG, Weixin *et al*. GPT Detectors are biased against non-native English Writers. **Patterns,** v. 4, n. 7, p. 1-4, 2023. DOI: https://doi.org/10.1016/j.patter.2023.100779.

LINDEBAUM, Dirk; FLEMING, Peter. ChatGPT undermines human reflexivity, scientific responsibility and responsible management research. **British Journal of Management**, v. 35, n. 2, p. 566-575, 2024. DOI: https://doi.org/10.1111/1467-8551.12781

LIAO, Q. Vera.; VAUGHAN, Jennifer Wortman. AI transparency in the age of LLMs: a human-centered research roadmap. **arXiv preprint** arXiv:2306.01941, p. 5368-5393, 2023. DOI: https://doi.org/10.48550/arXiv.2306.01941.

LIMONGI, Ricardo. The use of artificial intelligence in scientific research with integrity and ethics. **Future Studies Research Journal: Trends and Strategies**, v. 16, n. 1, p. e845-e845, 2024. DOI: https://doi.org/10.24023/FutureJournal/2175-5825/2024.v16i1.845.

LIMONGI, Ricardo; MARCOLIN, Carla Bonato. AI Literacy Research: Frontier for High-Impact Research and Ethics. **BAR-Brazilian Administration Review**, v. 21, n. 3, p. e240162, 2024. DOI: https://doi.org/10.1590/1807-7692bar2024240162

LOPEZOSA, Carlos. La Inteligencia Artificial en los procesos editoriales de las revistas académicas: propuestas prácticas. **Infonomy**, v. 1, n. 1, 2023. DOI: https://doi.org/10.3145/infonomy.23.009.

LU, Q., QIU, B., DING, L., XIE, L. & TAO, D. Error Analysis Prompting Enables Human-Like Translation Evaluation in Large Language Models. **arXiv preprint** arXiv:2303.13809, 2023. DOI: https://doi.org/10.48550/arXiv.2303.13809.

MANGOLD, Sarah; REAM, Margie. Artificial Intelligence in Graduate Medical Education Applications. **Journal of Graduate Medical Education**, v. 16, n. 2, p. 115-118, 2024. DOI: https://doi.org/10.4300/JGME-D-23-00510.1.

MENDONÇA, Ricardo Fabrino; ALMEIDA, Virgílio; FILGUEIRAS, Fernando. **Algorithmic institutionalism:** the changing rules of social and political life. Oxford University Press, 2024.

MOORHOUSE, Benjamin Luke; YEO, Marie Alina; WAN, Yuwei. Generative AI tools and assessment: Guidelines of the world's top-ranking universities. **Computers and Education Open**, v. 5, p. 1-10, 2023. DOI: https://doi.org/10.1016/j.caeo.2023.100151.

MOROZOV, Evgeny. **Big Tech:** a ascensão dos dados e a morte da política. São Paulo: Ubu, 2018.

NARAYANAN, Arvind; KAPOOR, Sayash. **AI snake oil:** what Artificial Intelligence can do, what it can't, and how to tell the difference. Princeton University Press, 2024.

NATURE. Why Nature will not allow the use of generative AI in images and videos. **Nature**, v. 618, 2023. Disponível em: https://doi.org/10.1038/d41586-023-01546-4.

NG, D. T. K. et al.. Conceptualizing AI literacy: An exploratory review. Computers and Education: Artificial Intelligence, v. 2, 100041, 2021. Disponível em: https://doi.org/10.1016/j.caeai.2021.100041

OXFORD. **Publication and authorship**. 2023. Disponível em: https://researchsupport.admin.ox.ac.uk/governance/integrity/publication.

OXFORD UNIVERSITY PRESS. **How are researchers responding to AI?** 2024. Disponível em: https://corp.oup.com/news/how-are-researchers-responding-to-ai/. Acesso em: 4 nov. 2024.

PEREIRA, Ricardo *et al*. Generative Artificial Intelligence and academic writing: an analysis of the perceptions of researchers in training. **Management Research**, v.

1, p. 1-10, 2024. DOI: https://doi.org/10.1108/MRJIAM-01-2024-1501.

PERES, Frederico. A literacia em saúde no ChatGPT: explorando o potencial de uso de inteligência artificial para a elaboração de textos acadêmicos. **Ciência & Saúde Coletiva**, v. 29, p. e02412023, 2024. DOI: https://doi.org/10.1590/1413-81232024291.02412023

PERKINS, Mike *et al*. Detection of GPT-4 generated text in Higher Education: combining academic judgement and software to identify generative AI tool misuse. **Journal of Academic Ethics**, v. 22, n. 1, p. 89-113, 2024. DOI: https://doi.org/10.1007/s10805-023-09492-6.

PRATHER, A. et al. The widening gap: the benefits and harms of Generative AI for novice programmers. **Arvix**, 28 maio 2024. Disponível em: https://arxiv.org/abs/2405.17739

RAHMAN, Mizanur *et al*. ChatGPT and academic research: a review and recommendations based on practical examples. **Journal of Education, Management and Development Studies**, v. 3, n. 1, p. 1-12, 2023. DOI: http://doi.org/10.52631/jemds.v3i1.175.

RAMOS Anatália Saraiva Martins. Generative Artificial Intelligence based on Large Language Models: tools for use in academic research. **SciELO Preprints**, 2023. DOI: https://doi.org/10.1590/SciELOPreprints.6105.

RAY, Partha Pratim. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. **Internet of Things and Cyber-Physical Systems**, v. 3, p. 121-154, 2023. DOI: https://doi.org/10.1016/j.iotcps.2023.04.003.

RESNIK, David B..; HOSSEINI, Mohammad. The ethics of using Artificial Intelligence in scientific research: new guidance needed for a new tool. **AI and Ethics**, p. 1-23, 2024. DOI: https://doi.org/10.1007/s43681-024-00493-8.

RIKAP, Cecilia. **Capitalism, power and innovation:** intellectual monopoly capitalism uncovered. Londres: Routledge, 2021.

RODRIGUES, Gisele da Silva; BRANDÃO, Valéria Ramos de Amorim; TRIVELATO, Rosana Matos da Silva. ChatGPT: uma ferramenta para a pesquisa científica. **Código 31: Revista de Informação, Comunicação e Interfaces**, v. 2, n. 1, 2024. DOI: https://doi.org/10.70493/cod31.v2i1.9916.

RÖHE, Anderson; SANTAELLA, Lucia. Confusões e dilemas da antropomorfização das inteligências artificiais. **TECCOGS: Revista Digital de Tecnologias Cognitivas**, n. 28, p. 67–75, 2023. DOI: https://doi.org/10.23925/1984-3585.2023i28p67-75.

SABBATINI, Marcelo. Do plágio à publicidade disfarçada: brechas da fraude e do antiético na comunicação científica. **ComCiência** (UNICAMP), v. 1, p. 3, 2013. Disponível em: https://comciencia.scielo.br/scielo.php?script=sci_arttext&pid=S1519-76542013000300009&lng=pt&nrm=iso.

SADASIVAN, V. S. et al. Can AI-Generated text be reliably detected? **arXiv**, art. arXiv:2303.11156v2, 2023. Disponível em: https://doi.org/10.48550/arXiv.2303.11156

SAMPAIO, Rafael Cardoso *et al*. ChatGPT e outras IAs transformarão a pesquisa científica: reflexões sobre seus usos. **Revista de Sociologia e Política**, v. 32, p. 1-24, 2024a. DOI: https://doi.org/10.1590/1678-98732432e008.

SAMPAIO, Rafael Cardoso *et al*. Uma revisão de escopo assistida por Inteligência Artificial (IA) sobre usos emergentes de ia na pesquisa qualitativa e suas considerações éticas. **Revista Pesquisa Qualitativa**, v. 12, p. 01-28, 2024b. DOI: https://doi.org/10.33361/RPQ.2024.v.12.n.30.729.

SAMUELSON, Pamela. Generative AI meets copyright. **Science**, v. 381, n. 6654, p. 158-161, 2023. DOI: https://doi.org/10.1126/science.adi0656.

SANTAELLA, Lucia. Por que é imprescindível um manual ético para a Inteligência Artificial Generativa? **TECCOGS: Revista Digital de Tecnologias Cognitivas**, n. 28, p. 7–24, 2023. DOI: https://doi.org/10.23925/1984-3585.2023i28p7-24.

SARAIVA Jr., Francisco. Transformando a sala de aula: utilizando a Inteligência Artificial generativa no aprendizado ativo. **Revista Brasileira de Casos de Ensino em Administração**, v. 14, n. especial1, p. a16-a16, 2024. DOI: http://dx.doi.org/10.12660/gvcasosv14nespecial1a16.

SCHULHOFF, Sander *et al*. The Prompt Report: a systematic survey of prompting techniques. **arXiv preprint** arXiv:2406.06608, 2024. DOI: https://doi.org/10.48550/arXiv.2406.06608.

SILVA, Maria Eduarda Ramos da; SERRANO, Paulo Henrique Souto Maior. Análise de sentimentos em textos de redes sociais: uma comparação entre o ChatGPT e métodos tradicionais. **Cadernos de Comunicação**, v. 27, n. 3, 2023. Disponível em: https://periodicos.ufsm.br/ccomunicacao/article/view/84828.

SILVA, Sivaldo Pereira da. Democracia, Inteligência Artificial e desafios regulatórios: direitos, dilemas e poder em sociedades datificadas. **E-LEGIS: Revista Eletrônica do Programa de Pós-Graduação da Câmara dos**

Deputados, v. 13, p. 226-248, 2020. DOI: https://doi.org/10.51206/e-legis.v13i33.600.

SILVA, Victo J.; BONACELLI, Maria Beatriz M.; PACHECO, Carlos A. Framing the effects of machine learning on science. **AI & Society**, v. 39, n. 2, p. 749–765, abr. 2024. DOI: https://doi.org/10.1007/s00146-022-01515-x.

SILVA, Tarcízio. **Racismo algorítmico:** inteligência artificial e discriminação nas redes digitais. São Paulo: Edições Sesc SP, 2022. Disponível em: https://racismo-algoritmico.pubpub.org/. Acesso em 6 nov. 2024.

SOARES, Bruno Johnson *et al.* Implicações da Inteligência Artificial na Educação. **TECCOGS – Revista Digital de Tecnologias Cognitivas,** n. 28, 2023, p. 76-86. DOI: https://doi.org/10.23925/1984-3585.2023i28p76-86.

SOCIEDADE BRASILEIRA DE COMPUTAÇÃO (SBC). **Plano de Inteligência Artificial da Sociedade Brasileira de Computação**. Porto Alegre: SBC, 2024. 19 p. DOI: https://doi.org/10.5753/sbc.rt.2024.141.

SOHAIL, Shahab Saquib *et al*. Decoding ChatGPT: a taxonomy of existing research, current challenges, and possible future directions. **Journal of King Saud University-Computer and Information Sciences**, p. 1-23, 2023. DOI: https://doi.org/10.1016/j.jksuci.2023.101675.

SPINAK, Ernesto. IA: Como detectar textos produzidos por chatbox e seus plágios. **Blog da SciELO**, 2023b. Disponível em: https://blog.scielo.org/blog/2023/11/17/ia-como-detectar-textos-produzidos-por-chatbox-e-seus-plagios/. Acesso em: 8 dez. 2023.

_____. Inteligência Artificial e a comunicação da pesquisa. **Blog da SciELO**. 2023a. Disponível em: https://blog.scielo.org/blog/2023/08/30/inteligencia-artificial-e-a-comunicacao-da-pesquisa/. Acesso em: 8 dez. 2023.

STAHL, Bernd Carsten; EKE, Damian. The ethics of ChatGPT: exploring the ethical issues of an emerging technology. **International Journal of Information Management,** v. 74, p. 1-14, 2024. DOI: https://doi.org/10.1016/j.ijinfomgt.2023.102700.

STOKEL-WALKER, Chris. ChatGPT listed as author on research papers: many scientists disapprove. **Nature,** vol. 613, no. 7945, p. 620-62, 2023. Disponível em: https://www.nature.com/articles/d41586-023-00107-z.

SUSARLA, Anjana *et al*. The Janus Effect of generative AI: charting the path for responsible conduct of scholarly activities in information systems. **Information Systems Research**, v. 34, n. 2, p. 399-408. DOI: https://doi.org/10.1287/isre.2023.ed.v34.n2.

TAYLOR & FRANCIS. **Taylor & Francis clarifies the responsible use of AI tools in academic Content Creation**, 2023. Disponível em: https://newsroom.taylorandfrancisgroup.com/taylor-francis-clarifies-the-responsible-use-of-ai-tools-in-aademic-content-creation/.

TEDESCO, A. L.; FERREIRA, J. L. Ética e Integridade acadêmica na Pós-Graduação em Educação em tempos de Inteligência Artificial. **Horizonte**s, v. 41, n. 1, p. e023032, 2023. Disponível em: https://revistahorizontes.usf.edu.br/horizontes/article/view/1620 .

THORP, H. Holden. ChatGPT is fun, but not an author. **Science***,* v. 379, n. 6630, p. 313-313, 2023. DOI: https://doi.org/10.1126/science.adg7879.

UNESCO. **Guia para a IA Generativa na educação e na pesquisa**. UNESCO Publishing, 2024. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000390241. Acesso em: 6 nov. 2024.

UNIÃO EUROPEIA. **Living guidelines on the responsible use of generative AI in research.** Bruxelas: European Commission, 2024. Disponível em: https://european-research-area.ec.europa.eu/news/living-guidelines-responsible-use-generative-ai-research-published. Acesso em: 23 out. 2024.

VELEZ, Meghan; RISTER, Alex. Beyond academic integrity: navigating institutional and disciplinary anxieties about AI-assisted authorship in technical and professional communication. **Journal of Business and Technical Communication,** p. 1-18, 2024. DOI: https://doi.org/10.1177/10506519241280646.

WILEY. **Authors and contributors. Best Practice Guidelines on Research Integrity and Publishing Ethics**. 2023. Disponível em: https://authorservices.wiley.com/ethics-guidelines/index.html. Acesso em: 6 nov. 2024.

WOHLIN, Claes. Guidelines for snowballing in systematic literature studies and a replication in Software Engineering. *In*: **Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering**, 2024. DOI: https://doi.org/10.1145/2601248.2601268 .

ZIELINSKI, Chris *et al*. **Chatbots, Generative AI, and scholarly manuscripts:** WAME recommendations on chatbots and Generative Artificial Intelligence in relation to scholarly publications. WAME. WAME, 31 maio 2023. Disponível em: https://wame.org/page3.php?id=106. Acesso em: 6 nov. 2024.

ZIEMS, Caleb *et al*. Can Large Language Models transform Computational Social Science? **Computational Linguistics**, v. 50, n. 1, p. 237-291, 2024. DOI: https://doi.org/10.1162/coli_a_00502

Apêndice 1: Ferramentas de IA citadas

Grandes Modelos de Linguagem

- **ChatGPT**: https://chatgpt.com/
- **Claude**: https://claude.ai/
- **Copilot**: https://copilot.microsoft.com/
- **Gemini**: https://gemini.google.com/
- **Hugging Face:** https://huggingface.co/
- **Llama:** https://www.llama.com/
- **Maritalk:** https://chat.maritaca.ai/
- **Mistral:** https://chat.mistral.ai/
- **Storm:** https://storm.genie.stanford.edu/

Modelos Generativos de imagem

- **Dall-e:** https://openai.com/index/dall-e-3/
- **Midjourney:** https://www.midjourney.com/home
- **Stable Diffusion:** https://stability.ai/stable-assistant

Modelos especializados em Programação

- **Codestral:** https://mistral.ai/news/codestral/
- **Code Llama:** https://www.llama.com/code-llama/
- **Github Copilot:** https://github.com/features/copilot
- **Star Coder**: https://huggingface.co/blog/starcoder

Outras ferramentas de IA

- **AvidNote:** https://avidnote.com/
- **Chatdoc:** https://chatdoc.com/
- **ChatPDF**: https://www.chatpdf.com/
- **Compose:** https://chromewebstore.google.com/detail/compose-ai-ai-powered-wri/ddlbpiadoechcolndfeaonajmngmhblj
- **Connected Papers**: https://www.connectedpapers.com/
- **Consensus**: https://consensus.app/
- **Cursor:** https://www.trycursor.com/
- **Data Squirrel:** https://datasquirrel.ai/
- **Elicit**: https://elicit.com/
- **Explain Paper:** https://www.explainpaper.com/
- **Glasp:** https://glasp.co/
- **Grammarly:** https://www.grammarly.com/
- **Harpa**: https://chromewebstore.google.com/detail/harpa-ai-automation-agent/
- **Heuristi.ca:** https://www.heuristi.ca/
- **Jenni:** https://jenni.ai/
- **Julius:** https://julius.ai/
- **Humata:** https://www.humata.ai/
- **Inciteful**: https://inciteful.xyz/
- **Iris:** https://iris.ai/
- **Kenenious.com:** https://keenious.com/
- **LanguageTool**: https://languagetool.org/pt/
- **Litmaps**: https://www.litmaps.com/
- **Merlin**: https://chromewebstore.google.com/detail/merlin-ask-ai-to-research/
- **MyReader**: https://www.myreader.ai/
- **Napkin:** https://www.napkin.ai/
- **Nested Knowledge:** https://nested-knowledge.com/
- **Notebook LM:** https://notebooklm.google.com/
- **Paperpal:** https://paperpal.com/
- **Perplexity**: https://www.perplexity.ai/
- **Quillbot**: https://quillbot.com/
- **Research Rabbit**: https://researchrabbitapp.com/
- **Rows:** https://rows.com/
- **Rytr**: https://rytr.me/
- **Sakana AI Scientist:** https://sakana.ai/ai-scientist/
- **Scholarcy**: https://www.scholarcy.com/
- **Scispace**: https://www.typeset.io/
- **SciSummary**: https://scisummary.com/
- **Scite**: https://scite.ai/
- **Smoodin:** https://smodin.io/
- **Text Cortex**: https://chromewebstore.google.com/detail/textcortex-assistente-pes/
- **Trinka:** https://www.trinka.ai/
- **Undermind:** https://undermind.ai/
- **Writefull**: https://writefull.com/
- **You**: https://you.com/

Listas de IAs acadêmicas

- Artificial Intelligence Applications for Social Science Research https://scholarsjunction.msstate.edu/ssrc-publications/6/
- Ferramentas de Inteligência Artificial para pesquisa científica https://osf.io/6dpfn/
- Generative A.I. Tools for Educators https://docs.google.com/spreadsheets/d/1SFPuZ0wh7v2RBuCTVmj9EUBSvLM7kZ4L-yLHG7e8Ygk/edit?gid=1952610479#gid=1952610479

Repositórios de IA

- **AI Library:** https://library.phygital.plus/
- **Futurepedia:** https://www.futurepedia.io/
- **Product Hunt:** https://www.producthunt.com/
- **There's An AI For That:** https://theresanaiforthat.com/
- **Top AI:** https://topai.tools/

Apêndice 2: Repositórios e dicas para prompts

Repositórios de prompts

- **Flow GPT:** https://flowgp.com/
- **GPT Prompt Tuner:** https://gptprompttuner.com/
- **IAEd Praxis:** https://iaedpraxis101.substack.com/s/prompts
- **Library - Anthropic:** https://docs.anthropic.com/en/prompt-library/library
- **OpenAI Platform:** https://platform.openai.com/docs/examples
- **Prompt Library:** https://www.moreusefulthings.com/prompts
- **Prompt Perfect:** https://promptperfect.jina.ai/
- **Snack Prompt:** https://snackprompt.com/

Materiais com dicas sobre a elaboração de prompts

- Guia da Anthropic com base no Claude: https://docs.anthropic.com/pt/docs/build-with-claude/prompt-engineering/overview
- Guia do Google tendo o Llama como exemplo: https://services.google.com/fh/files/misc/gemini-for-google-workspace-prompting-guide-101.pdf
- Guia da Meta para prompts tendo o Llama como exemplo: https://www.llama.com/docs/how-to-guides/prompting/
- Guia da Mistral: https://docs.mistral.ai/guides/prompting_capabilities/
- Guia da OpenAI com base no ChatGPT https://help.openai.com/en/articles/6654000-best-practices-for-prompt-engineering-with-the-openai-api

Outros guias gerais

- Rock Content: Guia completo de prompts para IA https://rockcontent.com/br/blog/guia-de-prompts-de-inteligencia-artificial-para-producao-de-conteudo/
- Dicas gerais para projetar prompts https://www.promptingguide.ai/pt/introduction/tips
- Effective Prompts for AI: The Essentials https://mitsloanedtech.mit.edu/ai/basics/effective-prompts/

Apêndice 3: Resumo dos princípios práticos para uso de IAG

Criamos com ajuda do “ChatGPT 4o com canvas” (uso em 2 de novembro de 2024), em uma aba protegida contra treinamento do modelo, um guia resumido para consultas mais rápidas. Os autores leram, corrigiram e incrementaram o resumo após a geração da IAG.

1 - Exploração inicial de ideias	Elaboração de temas, problemas e objetivos de pesquisa, teste inicial de hipóteses.
Oportunidades Facilitar o desenvolvimento de novas ideias, conexões criativas e identificação de pontos fortes e fracos da pesquisa.	**Cuidados** Priorizar a qualidade dos dados alimentados e prompts. Atentar para limitações e vieses dos resultados, buscando discernir conteúdo superficial. Priorizar fontes tradicionais e sua própria análise. Complementar o pensamento crítico com ajuda da IA, não o substituir.
2 - Busca de materiais acadêmicos	Processamento de grandes volumes de informações, identificação de padrões de relacionamento e artigos relevantes.
Oportunidades Apoiar os autores na realização de revisões de literatura, ajudando na pesquisa, classificação ou resumo de fontes bibliográficas. Como tal, fornece um ponto de partida, e não fim, para o exercício de revisão. Atuar como uma extensão de referências bibliográficas não encontradas nas buscas estruturadas em bases de dados como Scopus, Web of Science e afins.	**Cuidados** Utilizar ferramentas acadêmicas especializadas. Refinar buscas, através do uso consciente e refinado da ferramenta. Verificar qualidade das fontes, combinar com indexadores tradicionais (SciELO, Scopus, Web of Science etc.) e buscas abertas, Atentar para a cobertura de fontes em português e outros idiomas diferentes do inglês. Não usar LLMs de uso geral, sob risco de alucinação.
3 - Leitura e resumo de materiais acadêmicos	Auxílio para informações específicas, explicação de pontos complexos, resumo e síntese que auxiliam na decisão do que ler em profundidade Acesso à literatura em língua estrangeira..
Oportunidades Auxiliar na elaboração de resumos precisos de artigos acadêmicos, permitindo uma compreensão rápida e facilitando a seleção dos materiais que merecem leitura aprofundada.	**Cuidados** Evitar enviar materiais inéditos ou sensíveis. Validar os resumos gerados. Manter a leitura crítica dos materiais completos como passo essencial.
4 - Escrita	Geração de paráfrases, correções técnicas e melhorias no texto.

Oportunidades Automatizar tarefas relacionadas à edição e revisão do texto, aumentando a eficácia ao substituir processos antes terceirizados, como revisão gramatical e adequação de estilo.	**Cuidados** Seguir as diretrizes de transparência aplicáveis. Não “copiar e colar” textos gerados pela IA. Revisar e verificar as citações e a precisão acadêmica dos resultados. Manter sua própria voz autoral. Manter a revisão humana como elemento imprescindível.
5 - Análise e apresentação de resultados	Realização de análises quantitativas, qualitativas e geração de visualizações.
Oportunidades Facilitar a análise de grandes volumes de dados, automatizando a criação de visualizações, como gráficos e tabelas, e auxiliando na interpretação de resultados complexos. Direcionar a atenção dos autores para aspectos potencialmente interessantes de um conjunto de dados.	**Cuidados** Evitar compartilhar dados sensíveis de participantes. Manter ceticismo quanto aos resultados gerados. Validar usando métodos tradicionais. Documentar os passos. Refletir sobre transparência e replicabilidade dos resultados.
6 - Programação	Sugestão e correção códigos, facilitando melhorando a eficiência.
Oportunidades Gerar e corrigir códigos, sugerir melhorias e preparar documentação, permitindo um desenvolvimento mais eficiente de códigos de programação.	**Cuidados** Revisar o código gerado em termos de funcionalidade. Verificar eventuais falhas e vulnerabilidades no código, certificando-se da segurança. Evitar depender exclusivamente de IA.
7 - Transcrição	Aceleração significativa do processo de transcrição de áudio e vídeo.
Oportunidades Agilizar a transcrição de entrevistas e grupos focais, reduzindo o tempo gasto em atividades manuais. Aumentar a produtividade dos pesquisadores em estudos qualitativos.	**Cuidados** Não utilizar dados sensíveis e garantir privacidade dos dados dos participantes, especialmente voz (dado biométrico). Validar a precisão das transcrições, especialmente em gravações complexas com múltiplos falantes. Considerar elementos não-verbais importantes.
8 - Tradução	Apoio à leitura e estudo, redação de textos em línguas estrangeiras.
Oportunidades Facilitar a tradução de textos acadêmicos, especialmente para autores não nativos. Tornar o processo de publicação em revistas internacionais mais acessível e menos oneroso.	**Cuidados** Revisar cuidadosamente a tradução gerada. Buscar precisão e contextualização. Evitar traduções literais. Seguir diretrizes quanto ao uso de IA.
9 - Pareceres e avaliações	Realização de tarefas de avaliação, como verificação de formato, linguagem, qualidade da linguagem, detecção de plágio, entre outros.

Oportunidades Auxiliar a pré-avaliação de manuscritos, checando formatação, plágio, qualidade da linguagem e correspondência ao escopo da revista. Auxiliar o fluxo editorial.	**Cuidados** Não utilizar para artigos inéditos e dados sensíveis. Atentar para o viés em relação a métodos e variáveis. Utilizar preferencialmente como pré-avaliação, com sentido de apoio. Seguir diretrizes aplicáveis.
10 - Seleção	Elaboração de pré-projetos de pesquisa e simulação de defesas.
Oportunidades Auxiliar candidatos em processos seletivos, ajudando na elaboração de propostas e simulações de defesas, proporcionando um preparo mais completo e confiante. Ajudar na pré-avaliação de produtos textuais envolvidos no processo	**Cuidados** Avaliar textos considerando características de escrita humana. Não utilizar ferramentas de detecção de IA como única evidência de uso. Triangular avaliação de textos com entrevistas ou defesas para garantir a autenticidade do candidato. Manter a análise crítica por parte dos avaliadores.
11- Agentes de IA	Usos especializados em tarefas particulares, como exploração de bases de dados, realização de tarefas repetitivas e análises específicas.
Oportunidades Automatizar tarefas específicas e coordenadas, facilitando, por exemplo, a análise de bases de dados e a criação de fichamentos. Promover uma abordagem sistemática e eficaz de tarefas de pesquisa.	**Cuidados** Repetir cuidados relacionados à privacidade, ao viés, à possibilidade de erros e alucinações e à propriedade intelectual.
12 - Detecção de IA Generativa	Identificação de texto generativa para a redação de textos acadêmicos.
Oportunidades Auxiliar na garantia de transparência e na promoção de boas práticas acadêmicas. Educar autores sobre os limites do uso da IA e promover integridade nas publicações.	**Cuidados** Compreender que a precisão e confiabilidade destas ferramentas são contestadas. Entender que os valores fornecidos são estimativas, com alta margem de erro. Usar apenas como apoio e não como única evidência. Evitar abordagem puramente tecnológica, usando como suporte uma análise humana criteriosa.

Autores

Rafael Cardoso Sampaio

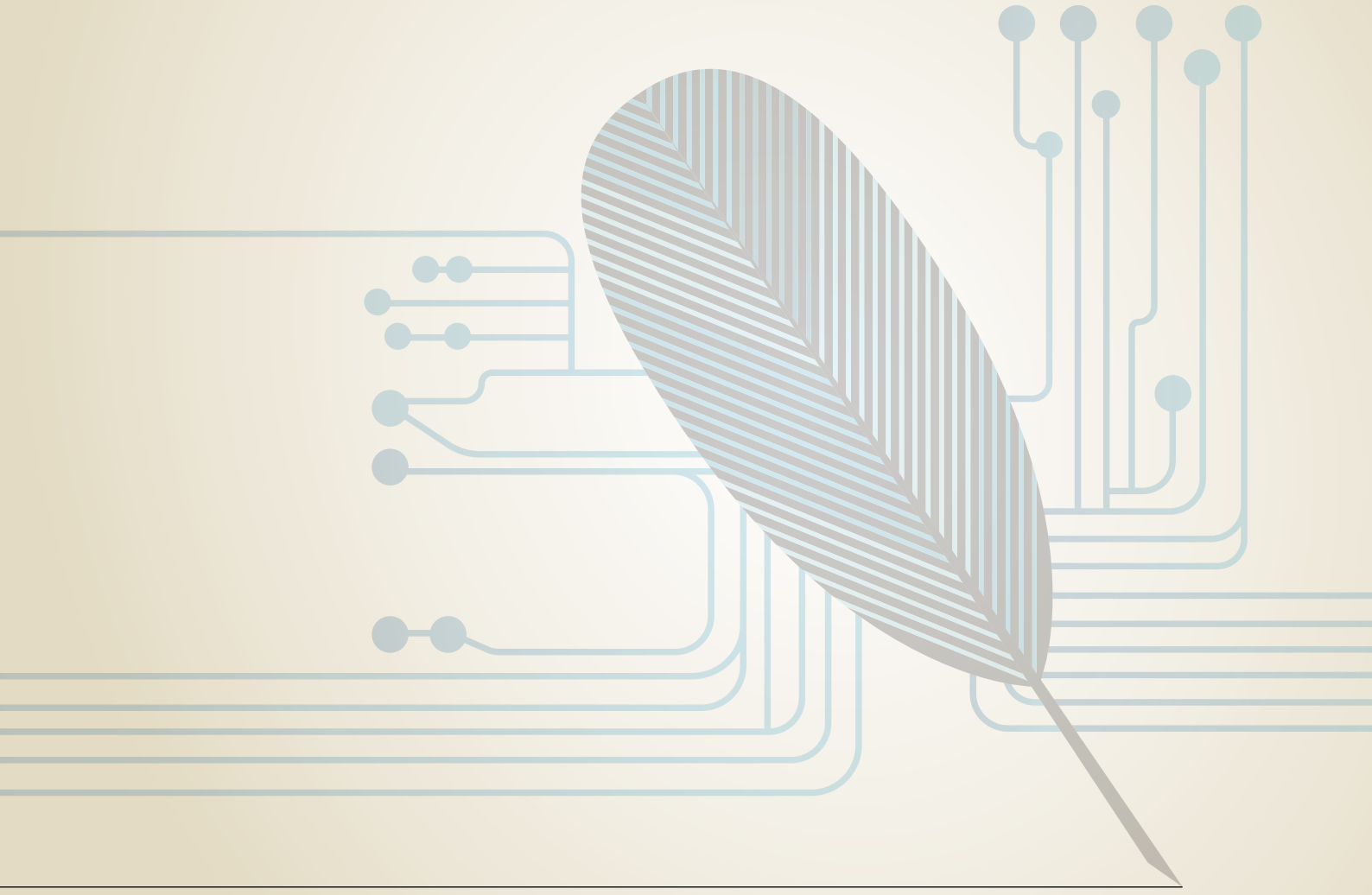
Professor permanente do Programa de Pós-graduação em Ciência Política (UFPR) e do Programa de Pós-graduação em Comunicação Social (UFPR). Pesquisador do Instituto Nacional de Ciência e Tecnologia em Democracia Digital (INCT-DD) e do Instituto Nacional de Ciência e Tecnologia Transformações da Participação, do Associativismo e do Confronto Político (INCT Participa). Coordenador do grupo de Pesquisa Comunicação Política e Democracia Digital (COMPADD). Bolsista produtividade do CNPq nível 1D. Foi presidente da Associação Brasileira de Pesquisadores em Comunicação e Política (Compolítica) entre 2019 e 2021. Atualmente, é diretor de pesquisa da Associação Brasileira de Ciência Política (ABCP). Desde 2008, pesquisa as relações entre tecnologias digitais e os regimes democráticos e agora tem se interessado no rigor das aplicações de métodos qualitativos e dos impactos da inteligência artificial generativa na pesquisa acadêmica. É autor, com Diógenes Lycarião, do livro "Análise de conteúdo categorial: manual de aplicação" publicado pela editora da ENAP. cardososampaio@gmail.com

Marcelo Sabbatini

Doutor em Teoria e História da Educação - Universidad de Salamanca (Espanha) em 2004. Mestre em Comunicação Social, modalidade Comunicação Científica e Tecnológica, Universidade Metodista de São Paulo, 2000. Atualmente é Professor Associado do Departamento de Fundamentos Sócio-Filosóficos da Educação do Centro de Educação da Universidade Federal de Pernambuco (UFPE). Professor-pesquisador do Programa de Pós-Graduação em Educação Matemática e Tecnológica - EDUMATEC-UFPE e membro do Grupo de Estudos em Novas Tecnologias e Educação (GENTE), vinculado ao CNPq. Associado da Intercom, onde foi vice, e logo coordenador do Grupo de Pesquisa em Folkcomunicação (2019-2024). Atualmente edita um boletim informativo com a temática da Inteligência Artificial na Educação: IAEdPraxis: Caminhos Inteligentes para a Educação (https://iaedpraxis101.substack.com/). marcelo.sabbatini@ufpe.br

Ricardo Limongi

Doutor em Administração na linha de Estratégias de Marketing pela FGV/SP, com estágio doutoral na Cornell University. Pós-doutorado em Economia Comportamental aplicada ao Marketing pela UnB e Pós-doutorado em Machine Learning aplicado ao Marketing pela UFRGS. Professor Associado na Universidade Federal de Goiás (UFG). Atualmente é Professor Visitante na Universidade Santiago do Chile (USACH) e foi na Mona University (UWIMONANA) na Jamaica. Professor Permanente no Programa de Pós-Graduação em Administração na UFG (Coordenador no Biênio 2020-2022) e na Universidade Federal de Uberlândia (UFU). Editor Chefe da Brazilian Administration Review (BAR) para o triênio 2024-2026, e editor associado na Revista de Administração de Empresas - RAE, na Revista de Administração Mackenzie - RAM (Qualis A2), PLOS One e Pretexto. Possui formação complementar em Estatística Espacial, Data Science e Machine Learning. Seus temas de interesse são: predição e business analytics; spatial data science; demand e estimação de mercado; inteligência artificial; machine learning. ricardolimongi@ufg.br



INTERCOM

ABCP
Associação Brasileira de Ciência Política

ANPOCS
Associação Nacional de Pós-Graduação e Pesquisa em Ciências Sociais