

inteligência artificial na pesquisa científica

Apresentação

Potencial e riscos da IA na ciência

Glauco Arbix

Este dossiê apresenta um painel em que especialistas debatem como percebem e analisam o potencial da inteligência artificial (IA) na pesquisa científica, sem camuflar os riscos que representa para a vida em sociedade. A USP, mais uma vez, marca sua presença ao avaliar a mais poderosa tecnologia que a humanidade já criou.

Os embates tecnológicos e científicos ocorrem em meio ao esforço para integrar cada vez mais a IA na vida econômica e social, o que suscita dúvidas e medo sobre o grau de controle que a humanidade pode e quer ter sobre as novas tecnologias. A promessa de ganhos de produtividade e de eficiência contrasta com resultados adversos, algumas vezes inesperados, no mercado de trabalho, nas desigualdades sociais, na democracia. Dilemas éticos e morais surgem a cada experimento e jus-

tificam a busca de um marco regulatório capaz de colocar limites sem asfixiar a criatividade e a inovação.

Como se sabe, a IA não é fruto apenas de *coders*, que mergulham no mar de dados e atravessam circuitos eletrônicos, fascinados que são pelo silêncio dos laboratórios. A busca do algoritmo perdido envolve tensões econômicas e sociais, que moldam comportamentos e alteram os parâmetros dos modelos estatísticos.

A IA é a mais poderosa tecnologia que a humanidade criou. É importante demais para que apenas poucas e gigantescas corporações tenham controle praticamente total sobre suas características e recursos. No mundo todo, o poder público tem a obrigação de colocar limites e salvaguardas para o uso,

GLAUCO ARBIX é professor titular de Sociologia da FFLCH-USP, pesquisador do Center for Artificial Intelligence USP-Fapesp-IBM e do Observatório da Inovação do Instituto de Estudos Avançados (IEA) da USP e ex-presidente do Ipea e da Finep.

a pesquisa e o desenvolvimento da IA. Para manter a sociedade protegida e não apenas correr atrás do fato consumado.

O ano de 2023 foi pródigo em multiplicar iniciativas voltadas para regulamentar a IA no plano global. O G7 em Hiroshima¹, a OCDE, a atuação da *Global Partnership on Artificial Intelligence* e as Nações Unidas², assim como a cúpula em Bletchley Park, no Reino Unido³, reafirmaram que a IA pode contribuir para resolver problemas históricos da humanidade e impulsionar um crescimento econômico inclusivo e sustentável, desde que utilizada de um modo seguro, confiável, responsável e com foco no ser humano.

Embora tenham sido raras as decisões de um aparato regulatório aprovado em forma de lei, com capacidade de efetivamente colocar limites e de avançar mecanismos de proteção à sociedade, vários governos com peso na geopolítica global começaram a reconhecer os riscos que uma IA desgovernada representa para nações e povos.

As tecnologias abrigadas debaixo do termo IA aumentaram recentemente seu poder de impacto nas mais variadas dimensões da atividade humana com a dissemi-

nação de sistemas generativos, baseados em *foundation models*, a exemplo do Chat-GPT. São imensas as oportunidades abertas capazes de transformar a realidade da vida cotidiana no planeta, a começar pela educação e saúde, moradia, segurança alimentar, pesquisa científica, ambiente e clima. Os riscos, porém, são cada vez mais presentes.

A COLETÂNEA DE ARTIGOS

O presente dossiê, um exercício que vai além da informação, sugere que a reflexão crítica que permeia a vida universitária é decisiva para o avanço de uma ciência que busca o bem comum. Ainda mais após um ano como 2023, marcado pela consolidação da IA como ponto de apoio fundamental para o avanço científico e tecnológico.

De acordo com a Agência Nacional de Ciência da Austrália, 99% dos campos de pesquisa global produzem resultados com base em técnicas de IA. A ascensão dos modelos generativos, após o lançamento público de sua versão 3.5 pela OpenAI, empresa sediada na Califórnia (Estados Unidos)⁴, representou um ponto de inflexão na trajetória da IA e abriu amplas avenidas para a pesquisa no mundo todo.

De um lado, esse movimento democratiza o acesso à atividade científica. O que antes requeria *expertise* para se lidar com a linguagem dos programadores, agora é feito com ferramentas amigáveis e de fácil

1 *G7 Action Plan for promoting global interoperability between tools for trustworthy AI* (Hiroshima, 20/5/2023).

2 O secretário-geral da ONU criou um *AI Advisory Body on risks, opportunities, and international governance of artificial intelligence*, com mais de 30 especialistas de alto nível dos mais diferentes países (Nova York, 27/10/2023).

3 28 países e a União Europeia se reuniram no Reino Unido e assinaram a *The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023*. Assinantes: Austrália, Brasil, Canadá, Chile, China, União Europeia, França, Alemanha, Índia, Indonésia, Irlanda, Israel, Itália, Japão, Quênia, Arábia Saudita, Holanda, Nigéria, Filipinas, Coreia, Ruanda, Singapura, Espanha, Suíça, Turquia, Ucrânia, EAU, Reino Unido e EUA.

4 A OpenAI foi criada em 2015 como entidade de pesquisa sem fins lucrativos. Em 2019 desenvolveu seu braço empresarial (com fins lucrativos), tornando-se uma instituição híbrida, que recebeu volumosos aportes de investimento da Microsoft.

acesso, com impactos positivos no universo do conhecimento e da sociedade. De outro, gera novos problemas e arrisca piorar o que já não era bom, como os empregos, a desigualdade e a própria democracia.

Mesmo com esses senões de luzes e sombras, a IA vem sendo abraçada por cientistas de todas as linhagens e tem se mostrado capaz de transformar o fazer científico e a atividade própria dos cientistas.

Para alguns expoentes da IA, como Demis Hassabis, neurocientista e fundador da DeepMind⁵, “estamos no renascimento da descoberta científica [...] e a IA tem um potencial incrível para ajudar a resolver os maiores desafios da humanidade”. A IA é um “multiplicador da engenhosidade humana”, completou Hassabis, que permite compreender e enxergar o que os cientistas não conseguem a olho nu⁶. Grande número de cientistas parece concordar com Hassabis e reconhecer o papel crescente que a IA desempenha nos mais diferentes campos da ciência.

Entre 1.600 cientistas entrevistados pela revista *Nature*⁷, mesmo sem deixar de registrar suas preocupações com os pontos negativos, cerca de 55% relataram se beneficiar da IA generativa para editar e traduzir textos do inglês; 40% utilizaram a IA para

acelerar e programar; mais de 30% para redigir artigos e para ampliar a pesquisa; 25% para obter ideias criativas; 15% para gerar novas hipóteses. A pesquisa registrou também que 58% dos cientistas consideram que a IA produz resultados não realizáveis por ferramentas tradicionais.

Para o Conselho Europeu de Pesquisa, em seu relatório de dezembro de 2023, 85% dos pesquisadores afirmaram que a IA generativa poderia assumir tarefas rotineiras e revisões de literatura, sendo que 40% consideraram que a IA aumenta significativamente a produtividade na ciência⁸.

O AVANÇO DA IA NA CIÊNCIA

A intensa penetração da IA nos sistemas de pesquisa e inovação pode ser percebida por alguns dos principais avanços que marcaram a ciência no ano de 2023 e que dificilmente seriam alcançados sem o concurso de tecnologias de IA.

No último dezembro, a FDA⁹ aprovou o primeiro e único medicamento baseado na tecnologia CRISPR¹⁰, o Casgevy, desenvolvido pela Vertex Pharmaceuticals (de Boston) e pela suíça CRISPR Therapeutics, capaz de tratar duas doenças: a anemia falciforme (doença sanguínea crô-

5 Empresa sediada em Londres cuja contribuição foi chave para a formulação das técnicas de aprendizagem profunda (*deep learning* – DL), baseada em redes neurais, que está na base da expansão da IA a partir de 2010. Ainda que a IA faça parte da caixa de ferramentas dos cientistas desde a década de 60, foi a partir da DL que abraçou o mundo.

6 Demis Hassabis, em palestra na University College London, em novembro de 2023, quando retomou a visão que tem sobre a IA desde entrevista à *The Economist*, em 2020.

7 Richard Van Noorden; Jeffrey M. Perkel. “AI and science: what 1,600 researchers think”. *Nature*, v. 621, 28/sep./2023.

8 *Foresight: use and impact of AI in the scientific process*. European Research Council, dec./2023.

9 Food and Drug Administration (FDA) é a agência pública norte-americana que controla alimentos, medicamentos, materiais biológicos e equipamentos de modo a garantir sua eficácia e segurança.

10 Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) é uma espécie de tesoura genética que permite cortes no código genético de uma célula, a partir da edição de seu DNA.

nica que afeta cerca de mais de 60 mil pessoas somente no Brasil, basicamente a população negra) e a β -talassemia (também genética e hereditária). Mais de 70% dos portadores de anemia falciforme vivem na África subsaariana.

Ainda no campo da saúde, a *Science* destacou o Ozempic e o Wegovy (nomes de mercado), produzidos pela Novo Nordisk (Dinamarca), para tratamento inovador de diabetes e obesidade e que também foram liberados pela FDA e autoridades europeias¹¹.

Também foram dignos de nota no ano os avanços promovidos pelo sistema AlphaFold2, desenvolvido pela Google DeepMind, que previu em 3D mais de 200 milhões de estruturas de proteínas e abriu espaço para a descoberta de novas drogas¹². E, na mesma direção, o destaque coube à primeira vacina contra a malária com mais de 80% de eficácia, desenvolvida pela Universidade de Oxford. Claro que avanços de ruptura sempre devem ser tratados com a devida cautela. A abordagem de novos métodos e medicamentos, assim como a discussão sobre a importância das novas tecnologias, fazem parte das preocupações do artigo “Inteligência artificial em saúde”, neste dossiê.

Novos instrumentos e metodologias para combater as mudanças do clima compõem os avanços científicos que ganharam destaque no ano passado. Ao contribuir para o enfrentamento da maior ameaça que atinge

a humanidade, a IA estimula a reação de todos os países para mitigar as consequências negativas das alterações climáticas. No médio e longo prazos, os modelos analíticos com base em algoritmos cada vez mais complexos serão cada vez mais fundamentais para as previsões do tempo e do clima. Os ganhos de precisão em modelos como os *earth system models* (ESM) permitem analisar as interações entre a atmosfera, oceano, ecossistemas terrestres, gelo e biosfera, como registrado no instigante artigo “Inteligência artificial e mudanças climáticas”.

No mesmo sentido, a busca de novas fontes de energia, de materiais avançados e novos experimentos se multiplica e anima redes globais de pesquisadores. No planejamento urbano, os profissionais de IA são cada vez mais requisitados para dar conta da complexidade da segurança e da mobilidade nas cidades do século XXI, que precisam se tornar mais inteligentes e sustentáveis, como aponta o artigo “Inteligência artificial, urbanização e cidades”, para melhorar a qualidade de vida de suas populações.

Na agricultura, uma das áreas mais importantes da economia brasileira, a IA está nos alimentos, nas fibras e na energia que movimenta empregos e racionaliza o uso do solo e da água, ao mesmo tempo em que ergue barreiras de defesa do meio ambiente. O artigo “A inteligência artificial na pesquisa agrícola” ilumina os avanços e os principais desafios que a agricultura enfrenta e que não podem prescindir das contribuições da IA.

UM ECOSSISTEMA PARA A IA

É preciso ir muito além das aplicações do conhecimento. Visão e governança precisam

11 Essas novas terapias não somente alteram como a obesidade e a diabetes são tratadas, mas também como são compreendidas. Porém, como toda abordagem de ruptura, outros efeitos são sentidos, nem sempre benéficos (*Science*, v. 382, issue 6676, 14/dec./2023).

12 Ver Database Protein Structure. Disponível em: <https://alphafold.ebi.ac.uk>.

ser reforçadas com a urgente formação de profissionais de IA e a construção de um sistema de dados capaz de valorizar os grandes ativos sediados no SUS, Inep, IBGE e tantos outros órgãos. A valorização de um ecossistema de IA consistente e diversificado ajuda a superar obstáculos crônicos (como o subfinanciamento) e a estimular a formação de uma comunidade de pesquisadores capaz de desenvolver uma IA inclusiva e sintonizada com a resolução dos problemas mais sentidos pela população brasileira.

Apenas um ecossistema nacional para IA, com suas regras e instituições, poderá abrir caminho para o equacionamento de questões éticas – como as dificuldades para explicar resultados científicos – ou para atuar no aumento da diversidade na pesquisa, que continua sendo marcada como um mundo masculino e branco. Passos dessa natureza ajudam a elevar a confiança na ciência e a preparar uma nova geração de pesquisadores, mais cosmopolitas, capaz de refletir sobre problemas locais, nacionais e globais. Sem abrir essa oportunidade, será difícil para o Brasil participar da criação de instituições globais de governança, aptas a definir os limites de uma IA e, assim, ajudar na construção da paz. Esse horizonte atrai jovens pesquisadores e ajuda a universidade a quebrar a atual estagnação do conhecimento que limita sua expansão.

IMPACTOS SOCIAIS DA IA

O Brasil não está na linha de frente da produção e desenvolvimento da IA. Como a maior parte dos países, o Brasil também tem dificuldades de infraestrutura computacional, de semicondutores, de acesso aos

bancos de dados que orientam os grandes modelos de linguagem existentes hoje. Dominados por poucas empresas, vivem o simulacro da autorregulação, que não tem efetividade e apenas prolonga a política que comandou a digitalização desde os anos 70 e que resultou no descontrole e na fragmentação atual, em que as sociedades são fustigadas pela desinformação massiva, que mina a busca da veracidade dos fatos, substrato básico da democracia.

Atualmente, tendências distintas estão em confronto e não contam com solução à vista. Uma destas caminha na direção de maior concentração e escala, procurando desenvolver sistemas cada vez maiores e mais poderosos. Ou seja, identificam melhor desempenho com maior poder computacional, mais dados e mais parâmetros. É o que movimenta os grandes modelos e responde por investimentos na casa de bilhões de dólares, difíceis de serem acompanhados pela maioria dos países e, claramente, por pequenas empresas. Por outro lado, sistemas de código aberto e menores procuram se aproximar das práticas de fronteira.

Seja qual for o desenlace desse confronto, a multimodalidade avançou rapidamente em 2023; ou seja, a capacidade da IA processar diferentes tipos de dados, não apenas textos, mas também imagens, vídeos, áudios e sons. Isso significa que os grandes modelos podem ser treinados em novos e vastos conjuntos de dados, oriundos de diferentes fontes, e que superam em muito as informações contidas apenas em textos. Essa tendência nem sempre ajuda os alinhamentos de valores, a possibilidade de sistemas de IA responderem em sintonia com os valores humanos, que, antes de tudo, são diferenciados e mutantes.

Esses processos procuram evitar respostas racistas e sexistas dos grandes modelos com

trabalho humano, com base em técnicas de *reinforcement learning*. Ou seja, avaliadores humanos identificam respostas inadequadas e treinam o sistema sobre o que é ou não aceitável. Como as variáveis e modulações são muitas, as falhas continuam a fazer parte do cotidiano dos modelos generativos, mesmo com todo o esforço de aprimoramento.

Pesquisadores do Allen Institute¹³ mostraram recentemente que a tentativa de aumentar o volume de dados e de utilizar o *reinforcement learning* para diminuir vieses pode resultar praticamente em seu contrário. A pesquisa testou cinco *chatbots* entre os mais populares, que responderam de modo diferente quando a pergunta no *prompt* era feita com a linguagem dos negros americanos ou em inglês padrão. Essa sensibilidade à linguagem diferente do inglês tido como correto expressou, segundo os pesquisadores, um racismo encoberto. Na verdade, os preconceitos que permeiam os bancos de dados e os algoritmos indicam a falta de diversidade no ambiente de pesquisa e nos grupos que desenvolvem a IA, como se pode acompanhar pela leitura do artigo “Elas na inteligência artificial – Questões de gênero”, tema que alimenta o debate sobre o florescimento de questões éticas, trabalhadas neste dossiê em um texto elegantemente construído com o título “Ética em IA e IA ética: prolegômenos e estudo de casos significativos”. O debate crescente em torno desse tópico apenas demonstra que as tecnologias disruptivas geram impactos que vão muito além das atividades de sua base tecnológica. No mesmo sentido, a circulação

intensa de conteúdos pelos canais da internet apenas reforça que as distorções da IA podem perpetuar vieses e injustiças sistêmicas. Como os dados e informações da rede global sustentam grande parte dos modelos generativos, o entrelaçamento do debate ético com a moderação de conteúdos torna-se inevitável. O trabalho incluído nesta coletânea “Inteligência artificial e internet: um olhar sobre o conteúdo de usuários e a sua moderação” ajuda a visualizar os riscos e o esforço de proteção que as plataformas de internet e os grandes modelos de linguagem precisam enfrentar. Essa perspectiva é ampliada pelo texto “Inteligência artificial, direito e pesquisa jurídica”, que procura mapear as implicações da institucionalização de diferentes projetos de IA no sistema de justiça brasileiro e as consequências para a pesquisa em direito.

A IA ENTRA NO DNA DA CIÊNCIA

A presente coletânea de artigos oferece uma pequena amostra de como a IA se tornou relevante e foi abraçada no mundo todo pela ciência e pelos pesquisadores, em um movimento que parece irreversível. Globalmente, o número de artigos em revistas científicas triplicou em uma década, de 2010 a 2021, enquanto os artigos baseados em *machine learning* quadruplicaram no mesmo período¹⁴.

A imprecisão, a segurança, as alucinações persistem como grandes obstáculos ao rápido crescimento da IA generativa nas economias. Mesmo assim, os dados são inequívocos: a partir de 2016, a IA saltou dos bancos

13 V. Hofmann et al. “Dialect prejudice predicts AI decisions about people’s character, employability, and criminality”, 2024. Disponível em: <https://arxiv.org/abs/2403.00742>.

14 Center for Security and Emerging Technology Report, 2022.

acadêmicos para os laboratórios de grandes corporações, que lideram desde então a geração de modelos generativos. Como *players* globais, mostram-se mais fortes que a maioria dos países. O controle que exercem sobre gigantescos bancos de dados e um parque de computadores de alto desempenho são uma amostra do nível de investimento que realizam para se reafirmarem como agentes globais determinantes do presente e provavelmente do futuro da IA. Esses grandes conglomerados expandem a IA ao mesmo tempo que constroem governos e inibem o esforço de construção de sistemas reguladores globais e nacionais.

Mais um paradoxo do mundo de hoje, em que o número de incidentes e controvérsias provocados pela IA cresceu 26 vezes entre 2012 e 2021, segundo o relatório da AI Index deste ano, publicado pela Universidade de Stanford. A opacidade das ferramentas da IA e suas consequências negativas, como a discriminação, as falsidades e a desinformação, são tratadas neste dossiê no texto “Desafios da IA responsável na pesquisa científica”.

O DECLÍNIO DA PESQUISA DE RUPTURA

A ciência precisa de uma IA transparente, capaz de ir além da lógica que preside a pesquisa há décadas, e que se apoia na revisão por pares, nos laboratórios universitários e nas declarações em favor da interdisciplinaridade.

O esforço de superação se justifica porque estudos recentes mostram uma situação incômoda para a ciência, ao apontar um arrefecimento do progresso da pesquisa nos grandes campos científicos. Nas palavras dos

pesquisadores, o número de artigos cresceu, mas não o de *papers* de ruptura, aqueles que abrem novas fronteiras para o conhecimento. Esse estudo foi publicado pela *Nature*, em 2023, e cobriu 45 milhões de artigos e 39 milhões de patentes, rastreados nos últimos 60 anos. Os resultados são fortes e pedem muita reflexão sobre as hipóteses que tentam explicar essa queda.

É possível que a ciência esteja com dificuldades para lidar com o enorme aumento do conhecimento acumulado; por isso, tenderia a apanhar os *low-hanging fruits*, as frutas mais maduras, os *insights* com origem nos principais autores e nas revistas mais importantes, tentando com isso trilhar os caminhos da pesquisa de menor risco. Uma segunda hipótese é que o crescimento da pesquisa não foi acompanhado do devido financiamento, ou seja, o subfinanciamento da ciência dificultaria os artigos de ruptura. A terceira hipótese, corrente nas universidades, é que os cientistas, pressionados e ávidos por publicar, tenderiam a diminuir a experimentação e sua ousadia.

A quarta hipótese, a que mais tem aderência à IA, sugere uma quase exaustão de recursos, metodologias e de teorias explicativas para novos fenômenos. O aprofundamento desta hipótese indica que o declínio das descobertas disruptivas pede mudanças fundamentais na natureza da produção científica e tecnológica.

A IA, claramente, tem potencial para ajudar em todas essas dimensões.

A TRAVESSIA ATUAL

Exatamente por isso, a IA se transformou em objeto de desejo de países,

empresas e universidades, que querem alavancar sua presença na arena mundial, nos mercados e nos domínios do saber. Em outras palavras, a IA passou a ocupar uma posição insubstituível para impulsionar todas as fases da pesquisa, seja na assistência, na automação, na previsão e na construção de cenários e estratégias de produção de conhecimento novo.

Apesar de amplamente disseminada pelo planeta, os destinos da inteligência artificial ainda não são nítidos. Todas as tensões presentes em sua criação, desenvolvimento e uso relatadas neste dossiê são expressão de uma disputa aberta entre países, empresas, universidades e diferentes grupos sociais. São dissonâncias de interesse, de hábito, de história, que diferenciam as intenções do gesto efetivo.

É importante lembrar que a IA está em sua primeira infância e suas características ainda estão em plena formação. Suas configurações aguardam definição, testes e provas de conceito. Ao se olhar para um futuro mais distante, combinações entre métodos simbólicos e baseados em dados provavelmente receberão maior atenção, enquanto apostas como computação quântica poderão levar a novos saltos de qualidade, como aponta o artigo desta coletânea “O futuro da pesquisa em inteligência artificial”.

Nossa época é de trânsito. E as novas tecnologias estão envelopadas pelo social, não o contrário. A proatividade é parte da ciência moderna, que se destacou como uma construtora de mundos, que recusou a passividade destilada pelas diferentes versões do determinismo tecnológico.

A ciência ousada recusa o deslumbremento e o medo diante do novo e do desconhecido. Por isso mesmo, o abraço sem a crítica e a crítica impotente da IA não conseguem explicar e muito menos modificar o curso, a configuração e os propósitos dessa tecnologia que altera padrões e sacode instituições e suas regras, a começar pela universidade e seus mecanismos de ensino, aprendizagem e pesquisa.

A certeza que se pode ter neste momento, diante de impactos diferenciados que a IA provoca na sociedade, é que nossas vidas serão muito diferentes. Em alguns domínios, como na ciência, a reviravolta deverá ser grande e duradoura, o que levará pesquisadores a abordar a IA com a prudência apropriada aos tempos de hoje. É o que está registrado neste dossiê: um tratamento tão atual quanto responsável da IA, temperado pela recomendação sensata do físico Mario Krenn, do Max Planck Institute: “Aproximem-se da IA como de uma musa, uma fonte de inspiração e de ideias”¹⁵.

Boa leitura.

15 Ver: <https://mariokrenn.wordpress.com/>.