



ARTIGO

PROMOÇÃO DA EVOLUÇÃO RESPONSÁVEL DE INTELIGÊNCIA ARTIFICIAL E TECNOLOGIAS CORRELATAS, MITIGANDO SEUS RISCOS E AMPLIANDO SEUS BENEFÍCIOS TECNOLÓGICOS, SOCIOECONÔMICOS E CULTURAIS

POR

André Carlos Ponce de Leon Ferreira de Carvalho, Claudia Bauzer Medeiros e Wagner Meira Jr.
andre@icmc.usp.br, cmbm@unicamp.br e meira@dcc.ufmg.br

A emergência de ferramentas de Inteligência Artificial (IA) [4] para apoiar decisões e soluções nos mais variados setores da sociedade, fruto de investimentos em pesquisa, desenvolvimento e inovação (PD&I) principalmente na área de Computação, trouxe vários benefícios socioeconômicos. Dada sua influência crescente em quase todas as áreas do conhecimento, e questões éticas associadas, novos avanços que possam estender esses benefícios de forma sustentável e inclusiva demandam ações coordenadas de PD&I com a participação

de todas as áreas da Computação.

O termo IA vem sendo usado como sinônimo de Grandes Modelos de Linguagem (LLMs, do inglês Large Language Models) [3], uma subárea de Redes Neurais Artificiais, por sua vez, uma das subáreas da IA. Os LLMs combinam algoritmos e técnicas sofisticadas de Processamento de Linguagem Natural e de Aprendizado de Máquina (AM). Não podemos esquecer que a IA é bem mais ampla, incluindo várias outras subáreas.

É possível considerar a IA como apoiada por três grandes grupos interligados - algoritmos que extraem conheci-

mento de dados, que, por demandas computacionais, precisam de plataformas de hardware e software poderosas. Para isso, requer conhecimentos e comportamento ético em gestão de dados, concepção de algoritmos e plataformas para sua execução. A IA cobre vários aspectos no uso e geração de dados e algoritmos, e sua execução, combinando aspectos práticos de implementação ao rigor teórico, além da interpretação dos resultados, que devem considerar o contexto de IA Responsável.

Em suma, os conceitos, fundamentos e algoritmos de IA têm sido, na maior parte, propostos e investigados por pesquisadores e profissionais da área da Computação. De forma complementar, todas as outras áreas do conhecimento também vêm contribuindo para esses avanços: ferramentas de IA estão sendo aplicadas e contribuem para o seu crescimento, acelerando o desenvolvimento de soluções, obtendo novos insights, apoiando a tomada de decisão e economizando tempo e trabalho manual. Muitos avanços em IA vêm sendo obtidos graças a esta colaboração e sinergia.

Dados, algoritmos e plataformas de software e hardware

Não é possível utilizar ferramentas de IA, ou fazer qualquer tipo de análise computacional de dados, sem pensar nos dados, estudados em Engenharia e em Ciência de Dados. Por outro lado, a capacidade da IA generativa de produzir grandes volumes de dados de forma não transparente (e eventualmente incorreta ou inconsistente) traz novos desafios em gestão e visualização de dados, revolu-

cionando o ciclo de vida de dados. Destacam-se, ainda, desafios associados à heterogeneidade, curadoria, proveniência e qualidade dos dados, além de propostas de novas linguagens que facilitem o acesso e a consulta a dados por usuários leigos.

A heterogeneidade nunca vai desaparecer: dados são sempre coletados e gerados por humanos, dispositivos ou de forma sintética, a partir de demandas e visões diferentes do mundo. Curadoria e limpeza são associadas à qualidade dos dados, que não é absoluta, mas depende do propósito para o qual são gerados. A curadoria de dados criados por IA é um desafio em aberto, envolvendo questões de proveniência e rastreabilidade. Novos desafios surgem quando dados são utilizados por algoritmos, modelos e sistemas. Um modelo de IA, em geral uma função matemática, é o cerne das aplicações de IA. Um sistema de IA é um conjunto de vários componentes, que inclui um ou mais modelos de IA, projetado para, de alguma forma, ser útil para os humanos. Surgem, assim, questões referentes ao projeto e implementação de algoritmos extensíveis e adaptáveis para a construção de ferramentas de IA e a adaptação e curadoria dos dados.

Atualmente, há um grande interesse em uma das subáreas da IA, a de AM, por conta dos avanços em aprendizado profundo e dos modelos como IA gerativa, que incluem os grandes modelos de linguagem (LLMs). Os LLMs tornaram-se possíveis com o desenvolvimento de algoritmos de aprendizado profundo (deep learning) para o treinamento de

redes neurais profundas [2], dando origem às redes neurais convolucionais e aos modelos transformadores (de transformers) [6].

Aqui, é preciso considerar confiabilidade, escalabilidade, explicabilidade e robustez dos algoritmos de IA desenvolvidos para aplicações gerais ou específicas. Para LLMs e modelos similares, é fundamental buscar maior controle e previsibilidade, para que tais modelos gerem resultados mais próximos dos esperados e especificados como, por exemplo, produzindo respostas com garantias de veracidade. Apesar desses avanços, para melhor usufruir dos benefícios da IA, é importante apoiar pesquisas em suas diferentes subáreas, incluindo seus fundamentos teóricos e matemáticos.

O terceiro grande grupo que influencia, e é influenciado por avanços em IA, é a concepção e desenvolvimento de plataformas de hardware e software eficientes, escaláveis e sustentáveis. Em hardware, o grande desafio é encontrar um bom compromisso entre alto desempenho, escalabilidade, baixo consumo de água e energia e baixa emissão de carbono no projeto de novos processadores (CPU e aceleradores) e redes de interconexão. É preciso ainda projetar arquiteturas eficientes para processar modelos com um número muito grande de parâmetros (> 1 trilhão). Em software, o grande desafio é desenvolver uma pilha de software que trate um número muito grande de agentes/processos/threads, permitindo projetar, orquestrar, escalonar e executar sistemas de IA em larga escala, de maneira eficiente. Em paralelo, é preciso considerar o reuso, a

manutenção e a extensibilidade das aplicações.

Visando a IA Responsável

Assim como em qualquer área da Computação, a IA não deve ignorar aspectos éticos, investigados na sua subárea IA Responsável, que combina questões computacionais, e de áreas como Filosofia e Psicologia. Nela, componentes e sistemas de IA incorporam princípios éticos no seu inteiro ciclo de vida: concepção, desenvolvimento, implantação, uso e monitoramento [5]. Isso inclui Beneficência (bem-estar, dignidade, sustentabilidade); Não-maleficência (privacidade, confiabilidade técnica); Justiça (não-discriminação, solidariedade); Autonomia (liberdade e poder de decidir); e Transparência (responsabilização, confiança, explicabilidade). Transformar esses princípios abstratos e subjetivos em mecanismos concretos é um desafio que cresce com a necessidade de considerar diferentes contextos, culturas e legislações, até em uma mesma instituição.

A promoção de uma IA ética deve considerar três desafios específicos. (1) o desenvolvimento de processos, ferramentas e técnicas que incorporem princípios éticos desde a concepção até a implementação; (2) o estabelecimento de fundamentos sólidos para a criação de normatizações e regulações que assegurem a conformidade com padrões éticos; e (3) a formação de profissionais e pesquisadores, fornecendo subsídios para programas educacionais que integrem ética, regulação e boas práticas ao ensino técnico e à pesquisa.

As estratégias para promover uma IA responsável envolvem a criação de ambientes e mecanismos formais que incentivem a cooperação multidisciplinar entre profissionais de Computação e de outras áreas do conhecimento, ampliando perspectivas e abordagens. Também é fundamental estabelecer requisitos éticos claros e aderentes aos princípios de IA responsável, incluindo métricas de análise, definição de limiares aceitáveis e processos de verificação e acreditação para garantir que os sistemas estejam alinhados a padrões éticos.

A IA e tecnologias correlatas são cada vez mais usadas para resolver um grande leque de questões socioeconômicas rele-

vantes, em todos os domínios de conhecimento. Exemplos incluem mitigar efeitos de migrações humanas via planejamento urbano, ou mudanças climáticas, ou ainda agricultura digital e produção e distribuição de alimentos – e, portanto, efeitos em saúde e qualidade de vida. Isto requer enfoques multidisciplinares e a colaboração de pesquisadores e técnicos em Computação com os especialistas de domínio que vão interpretar os resultados, ao mesmo tempo buscando o uso responsável e ético da IA. Finalmente, é necessário intensificar a qualificação de recursos humanos em IA e tecnologias correlatas, ainda insuficientes para a enorme demanda do Brasil [1].'

Referências

1. Almeida, V. A. F. et al. Recomendações para o Avanço da IA no Brasil da Academia Brasileira de Ciências. GT-IA da Academia Brasileira de Ciências. Academia Brasileira de Ciências, ISBN 97865-981763-0-3, <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> , 2023
2. Bengio, Y., Goodfellow, I. e Courville, A. Deep Learning. Massachusetts: MIT Press, EUA, 2017
3. Brown, T. B., Mann, B., Ryder, N., Subbiah M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. . Language models are few-shot learners. In Advances in Neural Information Processing Systems (NeurIPS), 2020.
4. Russel, S. E Norvig, P. Artificial Intelligence: A Modern Approach (4th. ed.). Prentice Hall Press, USA, 2020.
5. UNESCO. Recommendations on the Ethics of Artificial Intelligence. SHS/BIO/PI/2021/1, <https://unesdoc.unesco.org/ark:/48223/pf0000381137> , 2021
6. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., e Polosukhin, I. Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS), 2017



ANDRÉ CARLOS PONCE DE LEON FERREIRA DE CARVALHO é Professor Titular e atual Diretor do Instituto de Matemática e Ciência da Computação da Universidade de São Paulo (ICMC-USP). Atua nas áreas de Inteligência Artificial, Ciência de Dados e Aprendizado de Máquina. É Diretor do Centro Brasileiro de Pesquisa em IA Aplicada (IARA) e Coordenador do Instituto Nacional de Ciência e Tecnologia (INCT) IA para o Bem Social. Foi Vice-Presidente da Sociedade Brasileira de Computação (SBC) por dois mandatos, 2019-2021 e 2021-2023. Faz parte do comitê de especialistas do International AI Safety Report.



CLAUDIA BAUZER MEDEIROS é Professora Titular do Instituto de Computação da Unicamp. Desenvolve pesquisas em gerenciamento de dados científicos, considerando desafios associados à heterogeneidade, volume e complexidade. Ex-presidente da SBC (2003-2007), é uma das pioneiras brasileiras em Ciência Aberta. Premiada nacional e internacionalmente por atividades de ensino, pesquisa e iniciativas para atrair mulheres para a Computação. Comendadora da Ordem Nacional do Mérito Científico do Brasil; membro da Academia Brasileira de Ciências e da World Academy of Sciences.



WAGNER MEIRA JR. é Professor Titular de Ciência da Computação na Universidade Federal de Minas Gerais. É vice coordenador do Instituto Nacional de Ciência Tecnologia em Inteligência Artificial (IAIA) e foi diretor de Cooperação com Sociedades Científicas da SBC entre 2019 e 2023. Comendador da Ordem Nacional do Mérito Científico do Brasil. Atua na área de sistemas paralelos e distribuídos em larga escala, mineração de dados, aprendizado de máquina e suas aplicações tendo publicado mais de 300 artigos científicos em congressos e periódicos.