

CONTRIBUIÇÃO DA INTELIGÊNCIA ARTIFICIAL PARA PRÁTICAS DE CIÊNCIA ABERTA

João Ricardo de Melo, Edson A. Oliveira Junior (Orientador). E-mail:edson@din.uem.br

Universidade Estadual de Maringá, Centro de Tecnologia Maringá, PR.

Área e subárea do conhecimento conforme tabela do CNPQ/CAPES: Ciência da Computação/Computabilidade e Modelos de Computação.

Palavras Chave: Inteligência Artificial, Ciência Aberta, Reprodutibilidade, FAIR

RESUMO

Este estudo apresenta uma Revisão Sistemática da Literatura (RSL) sobre a contribuição da Inteligência Artificial para práticas de Ciência Aberta em Computação, com foco em aspectos como princípios FAIR, reprodutibilidade, metadados e software de pesquisa. A busca foi realizada na base ScienceDirect, resultando em 14 estudos selecionados que evidenciam como técnicas de aprendizado de máquina e processamento de linguagem natural podem automatizar a curadoria de dados, enriquecer descrições, avaliar conformidade e apoiar ambientes reprodutíveis. Os resultados mostram que a IA atua como facilitadora da adoção da ciência aberta, reduzindo barreiras práticas, elevando a qualidade e a confiabilidade dos registros científicos. Conclui-se que a integração entre IA e ciência aberta fortalece a transparência, a auditabilidade e a democratização da produção científica, desde que orientada por padrões éticos e de governança.

INTRODUÇÃO

A Ciência Aberta pode ser entendida como um conjunto de práticas que visam tornar a pesquisa científica mais acessível, transparente e colaborativa, abrangendo a disponibilização de dados, métodos, publicações, softwares e demais artefatos de pesquisa de forma livre e reutilizável (UNESCO, 2022). Nas últimas décadas, esse paradigma consolidou-se como essencial para a produção e disseminação do conhecimento, promovendo reprodutibilidade, democratização do acesso e maior impacto social das descobertas científicas (NATIONAL ACADEMIES OF SCIENCES, ENGINEERING, AND MEDICINE, 2018). Movimentos internacionais e iniciativas apoiadas por grandes programas de financiamento, como o *Horizon 2020* na Europa, reforçam a adoção de políticas e infraestruturas que sustentem a abertura da ciência em escala global.

Paralelamente, a Inteligência Artificial (IA) emergiu como área transformadora, com técnicas de aprendizado de máquina, mineração de dados e processamento de

linguagem natural, oferecendo soluções inovadoras para lidar com o volume crescente de informações científicas. Publicações recentes da ACM têm investigado os avanços na “abertura” da IA, discutindo como este conceito vem sendo influenciado por práticas de software *de código aberto* seus impactos na comunidade acadêmica.

O movimento de ciência aberta está provando ser extremamente eficaz, especialmente quando oferece acesso a código e dados: um estudo com 30 artigos de IA muito citados constatou que a reprodutibilidade é significativamente maior quando o código e os dados são compartilhados, destacando a importância da abertura para a replicabilidade na IA.

No contexto da publicação científica, a IEEE tem desenvolvido várias frentes de apoio à ciência aberta, desde o renovado foco em acesso aberto com o IEEE Access até iniciativas de orientação para autores que utilizam conteúdo gerado por IA e objetos que ampliam a visibilidade e as descobertas das pesquisas IEEE OpenIEEE Xplore (IEEE STANDARDS ASSOCIATION, 2020). A partir de uma revisão sistemática da literatura, foi possível ver tendências, identificar falhas e caracterizar as contribuições da IA para a abertura, confiabilidade e maturidade das pesquisas científicas.

MATERIAIS E MÉTODOS

Este estudo foi conduzido sob a forma de uma **Revisão Sistemática da Literatura (RSL)**.

Processo de busca

A revisão concentrou-se exclusivamente na base **ScienceDirect**, por sua ampla cobertura de periódicos relevantes na área de ciência da computação e ciência aberta. As buscas foram realizadas utilizando combinação de termos relacionados à Inteligência Artificial, ciência aberta e práticas FAIR.

A **string de busca principal** foi: ("artificial intelligence" OR "machine learning") AND ("open science" OR FAIR OR reproducibility OR "research software" OR metadata)

Essa consulta foi aplicada ao campo **Title, Abstract, Keywords** da ScienceDirect.

Critérios de inclusão: (i) Estudos que apresentassem aplicações explícitas de IA para práticas de ciência aberta em computação ou áreas computacionais afins; e (ii) Trabalhos que abordassem diretamente aspectos como FAIR (*Encontráveis, Acessíveis, Interoperáveis, Reutilizáveis*), reprodutibilidade, metadados, software de pesquisa ou ambientes reprodutíveis.

Critérios de exclusão: (i) Artigos sem ligação clara entre IA e práticas de ciência aberta; e (ii) Estudos puramente conceituais de ética em IA sem conexão direta com princípios FAIR ou reprodutibilidade.

Processo de seleção

A triagem inicial foi feita por títulos e resumos, seguida de leitura focal em seções de métodos e contribuições. Em seguida, foram extraídos: objetivos, técnicas de IA aplicadas, prática de ciência aberta apoiada, principais resultados e limitações.

O conjunto final resultou em **14 estudos**, que cobrem desde extração automática de metadados até frameworks de reprodutibilidade e aproximações entre FAIR e aprendizado de máquina. A tabela a seguir apresenta os estudos selecionados.

Tabela 1 - Estudos Filtrados na Base de Dados

Estudo	Ano de Publicação
An automated solution for measuring the progress toward FAIR	2022
The role of metadata in reproducible computational research	2021
dtolAI: Reproducibility for Deep Learning	2021
Taking a fresh look at FAIR for research software	2020
Computing environments for reproducibility: Capturing the whole environment	2019
A survey of reproducible research with a focus on high performance computing	2018
Towards a framework for interoperability and reproducibility of ML models under FAIR	2023
Towards Machine FAIR	2021
Automated metadata extraction using neural language processing	2020
Deep Learning based extraction of algorithmic metadata in full text scholarly articles	2022
BRYT: Automated keyword extraction for open datasets	2021
A local platform for user friendly FAIR data management and reproducible analytics	2022
Formal definition and implementation of reproducibility tenets for computational workflows	2020
AI for science and global citizens	2023

Fonte: Elaboração Própria

RESULTADOS E DISCUSSÃO

Os resultados desta revisão permitem observar que a Inteligência Artificial atua como um **agente transversal** em diferentes práticas de ciência aberta em computação, indo além da simples automação de tarefas técnicas. As evidências identificadas sugerem múltiplas camadas de impacto, que podem ser discutidas sob quatro dimensões: eficiência operacional, qualidade da informação, integração com ecossistemas existentes e implicações sociotécnicas.

Eficiência operacional e redução de barreiras

Um dos achados centrais é a capacidade da IA de reduzir barreiras práticas para adoção da ciência aberta. Processos historicamente custosos, como a anotação de metadados, a verificação de conformidade com padrões FAIR ou o rastreamento de ambientes de execução, passam a ser automatizados ou semiautomatizados. Isso amplia a adesão de pesquisadores que, em condições manuais, percebem práticas abertas como onerosas. Nesse sentido, a IA funciona como um “facilitador”, integrando práticas de abertura ao fluxo cotidiano de pesquisa, sem exigir esforço adicional.

Qualidade e confiabilidade da informação

Outro aspecto relevante é que o uso de NLP (Processamento de Linguagem Natural) e aprendizado profundo não apenas reduz custos, mas também **eleva a qualidade dos registros científicos**. Estudos mostram que a extração automática de termos e descrições produz metadados mais consistentes e menos sujeitos a omissões ou ambiguidades. Essa melhoria de qualidade é crucial para garantir **reprodutibilidade**

e reutilização confiáveis. Contudo, permanece o desafio de validar sistematicamente a precisão dos metadados gerados.

Implicações sociotécnicas e democratização

Um ponto transversal nos estudos é a **dimensão sociotécnica da ciência aberta**. O uso de IA não apenas facilita práticas técnicas, mas também promove democratização ao tornar dados e software mais acessíveis e compreensíveis. Por outro lado, surge a preocupação com a **dependência de soluções algorítmicas proprietárias**, que pode gerar novas formas de assimetria e opacidade — em contraste com os objetivos de abertura e transparência. Assim, torna-se essencial que os próprios sistemas de IA usados para ciência aberta sigam princípios FAIR, garantindo auditabilidade e reuso.

Desafios persistentes e lacunas

Apesar das contribuições claras, a revisão destaca desafios ainda em aberto:

- **Padronização insuficiente:** faltam consensos amplos sobre esquemas de metadados para modelos de aprendizado de máquina, dificultando a interoperabilidade entre plataformas.
- **Avaliação limitada:** poucas iniciativas comparam diferentes técnicas de extração automática em múltiplos domínios, o que limita soluções generalistas.
- **Integração em MLOps:** ainda não existem pipelines consolidados que incorporem avaliação FAIR e verificadores de reprodutibilidade como parte nativa do ciclo de desenvolvimento de modelos.
- **Ética e governança:** embora IA acelere a abertura, a governança de dados sensíveis (por exemplo, privacidade em datasets biomédicos) continua sendo um ponto crítico que não pode ser resolvido apenas por soluções técnicas.

CONCLUSÕES

A revisão mostrou que a Inteligência Artificial e a Ciência Aberta podem atuar de forma complementar, automatizando processos, ampliando a reprodutibilidade e fortalecendo a transparência da pesquisa. Quando integradas de maneira ética e validadas, contribuem para uma ciência mais confiável, acessível e democrática.

REFERÊNCIAS

UNESCO. **UNESCO Recommendation on Open Science**. Paris: United Nations Educational, Scientific and Cultural Organization, 2022. Disponível em: <https://www.unesco.org/en/open-science/about>

IEEE STANDARDS ASSOCIATION. IEEE 7010: A New Standard for Assessing the Well-being Implications of Artificial Intelligence. IEEE Transactions on Artificial Intelligence, 2020. Disponível em: <https://arxiv.org/abs/2005.06620>. Acesso em: 28 ago. 2025.

NATIONAL ACADEMIES OF SCIENCES, ENGINEERING, AND MEDICINE. Open Science by Design: Realizing a Vision for 21st Century Research. Washington, DC: The National Academies Press, 2018. ISBN 978-0-309-47624-9. DOI: <https://doi.org/10.17226/25116>.