

## Produção científica sobre Inteligência Artificial e ética: Uma análise bibliométrica

Sheilla da Silva Barroso<sup>1</sup>  ; Laurides Batista Cruz<sup>1</sup>  ; Daniele Carvalho Castro<sup>1</sup> 

<sup>1</sup>Escola de Saúde Pública de Sergipe, Fundação Estadual de Saúde, Revista Sergipana de Saúde Pública – Aracaju (SE), Brasil.

\*Autora correspondente: **Sheilla**: sheillasb@gmail.com

**Resumo:** A intersecção entre Inteligência Artificial (IA) e ética é um campo de estudo que busca abordar as implicações morais, sociais e jurídicas. Neste contexto, o objetivo é analisar a evolução das publicações científicas realizadas sobre a utilização de ferramentas de inteligência artificial e ética. Quanto ao tipo de estudo realizado, trata-se de pesquisa bibliométrica, com abordagem quantitativa e descritiva. A coleta de dados ocorreu entre 2018 e 2013, utilizando os descritores: “inteligência artificial e ética”. Os resultados demonstraram que em 2023 o número de publicações aumentou de 6 para 12. Conclui-se, então, que há escassez de publicações nos anos analisados que abordem o tema IA e ética.

**Palavras-chave:** inteligência artificial; ciência; ética.

### Artificial Intelligence and ethics: A bibliometric analysis

**Abstract:** The intersection between Artificial Intelligence (AI) and ethics is a field of study that seeks to address the moral, social and legal implications. In this context, the objective is to analyze the evolution of scientific publications made on the use of artificial intelligence and ethics tools. Regarding the type of study carried out, it is bibliometric research, with a quantitative and descriptive approach. The results demonstrated an increase in the number of articles published from 2019 to 2023. It is concluded that discussions on the topic of ethics and AI will still go through several developments, thus, it is suggested a future perspective to carry out studies with meta-analysis, expanding the time period of analysis in order to evaluate the frequency of publications on the topic.

**Keywords:** artificial intelligence; science; ethic.

### Inteligencia artificial y ética: Un análisis bibliométrico

**Resumen:** La intersección entre Inteligencia Artificial (IA) y ética es un campo de estudio que busca abordar las implicaciones morales, sociales y legales. En este contexto, el objetivo es analizar la evolución de las publicaciones científicas realizadas sobre el uso de la inteligencia artificial y herramientas éticas. En cuanto al tipo de estudio realizado, se trata de una investigación bibliométrica, con enfoque cuantitativo y descriptivo. La recolección de datos se realizó entre 2018 y 2013, utilizando los descriptores: “inteligencia artificial, ética”. Los resultados demostraron que en 2023 el número de publicaciones aumentó de 6 a 12. Se concluye que existe escasez de publicaciones en los años analizados que aborden el tema de IA y ética.

**Palabras clave:** inteligencia artificial; ciencia; principio moral.



## 1 Introdução

A inteligência artificial (IA) é uma conquista tecnológica revolucionária, que abre oportunidades inexploradas, na qual as tarefas complexas e as tomadas de decisões são baseadas em dados e algoritmos (De Felice *et al.*, 2022). Os avanços na IA penetraram em todas as esferas da sociedade, desde educação, saúde, indústria e veículos não tripulados. Acredita-se que a IA não é humanizada e que esta, por sua vez, trará desigualdade à sociedade (Evgrafova *et al.*, 2022).

A IA não é um conceito novo, porém indefinido, por abranger várias áreas do conhecimento (De Felice *et al.*, 2022). A popularização da IA iniciou nas décadas de 1980 e 1990, com o advento das redes neurais, sem muito sucesso devido à deficiência das Unidades de Processamento Gráfico (UPG). Uma IA confiável baseia-se em sete requisitos técnicos (agência humana e supervisão; robustez e segurança; privacidade e governança de dados; transparência; diversidade, não discriminação e justiça; bem-estar social e ambiental; e responsabilização) sustentados em três pilares principais: (1) legal; (2) ético; e (3) robusto (Díaz-Rodríguez *et al.*, 2023). Conforme Xue e Pang (2002), a IA está cada vez mais integrada na descoberta científica, auxiliando cientistas a gerar hipóteses, interpretar um conjunto de dados, e a compreender as aplicações, dentro da ciência fundamental baseada, promovendo o desenvolvimento contínuo e eficiente dos métodos científicos (Wang *et al.*, 2023).

No entanto, existem preocupações de que as decisões mediadas pela IA possam ser difíceis de validar ou possam resultar em decisões fraudulentas, desta forma, a interseção entre IA e ética é um campo de estudo relevante que busca abordar as implicações morais, sociais e legais. Outro ponto que traz discussão sobre o tema é a interação IA e ética, que levanta uma série de perguntas e desafios importantes que devem ser cuidadosamente considerados para garantir que a IA seja desenvolvida e utilizada de maneira responsável, justa e benéfica para a humanidade. Neste contexto, o objetivo deste artigo é analisar a evolução das publicações científicas realizadas sobre o uso da inteligência artificial e ética, tendo como outros objetivos adjacentes: (i) avaliar a qualidade das publicações científicas das IA no contexto ético; (ii) analisar os principais riscos e benefícios das IA.

## 2 Metodologia

Em relação ao tipo de estudo realizado, trata-se de uma pesquisa bibliométrica. O método bibliométrico é uma técnica quantitativa de medição dos índices de produção e disseminação do conhecimento científico (Araújo, 2006), na análise objetiva da produção científica sobre a inteligência artificial e a ética.

Para o desenvolvimento do estudo, estabeleceram-se quatro etapas: (i) fundamentação teórica da pesquisa, (ii) objetivo da pesquisa, (iii) protocolo de pesquisa, (iv) análise e interpretação dos resultados.

A fundamentação teórica da pesquisa e o objetivo, apresentados na introdução, compuseram a primeira etapa deste estudo bibliométrico. Na segunda etapa, denominada de protocolo de pesquisa, foram definidas a base de dados, as palavras-chave e a estratégia de busca (Figueredo *et al.*, 2020). A base de dados *Science Direct* (<http://www.sciencedirect.com>) foi escolhida por abranger os documentos em diferentes áreas científica, tecnológica e da saúde. As palavras-chave no idioma inglês: *artificial intelligence*, *ethics* foram definidas intencionadas ao objeto de estudo, fazendo o cruzamento com o operador booleano *AND*. A estratégia de busca utilizada foi: *artificial intelligence AND ethics*.

Na terceira etapa, coleta dos dados, foram incluídos artigos originais indexados no período entre 2018 e 2023. O levantamento dos dados ocorreu em setembro de 2023; tendo como critérios de inclusão todas as produções que abordassem o tema proposto no título, no idioma inglês, que estivessem disponíveis na íntegra, e on-line. Excluindo-se, portanto, teses, dissertações, artigos incompletos ou resenhas, editoriais e cartas.

Durante a quarta etapa, para análise dos dados, foi desenvolvido um banco de dados no *Microsoft Office Excel*. Em seguida, os resultados de pesquisa foram extraídos da base de dados no formato de arquivo *BibTeX* e importados para o *software Zotero*.

Neste programa, os artigos pesquisados foram organizados em “coleções”. A análise, a partir deste momento, passou a ser feita no *Zotero* que, reunindo os 3104 artigos, oriundos da base de dados *Science direct*, permitiu um filtro em que se restringiu a análise para documentos que contivessem em seu título os termos “*artificial intelligence*” and “*ethics*”. Para a disposição e apresentação dos resultados, última etapa do método bibliométrico, optou-se por utilizar os gráficos criados pelo *Excel*.

Por se tratar de um estudo do tipo bibliométrico, com a utilização de dados secundários, disponíveis gratuitamente em repositório de dados, não houve necessidade de submissão ao Comitê de Ética em Pesquisas (CEP).

### 3 Resultados e discussão

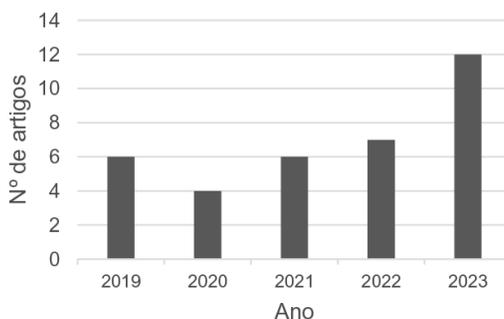
Para este estudo, foram identificados 3104 artigos, que após passar pelos critérios de inclusão adotados no estudo, resultaram em 35 artigos, posteriormente analisados. O demonstrativo da quantidade de artigos publicados relacionados ao tema objeto de estudo deste trabalho é representado pelo Gráfico 1. Observa-se um aumento no número de artigos publicados entre 2019 e 2023, com um aumento de 6 para 12 artigos publicados. O Gráfico 2 apresenta o número de publicações distribuída

por periódicos internacionais com o termo inteligência artificial e ética, no período de 2018-2023. O interesse pela IA pode ser demonstrado pelas recentes pesquisas que utilizam as aplicações relacionadas à IA no campo da indústria, tecnologia e da saúde (Pan, 2016; Dave; Patel, 2023).

De acordo com Pan (2016), com a popularização da internet, a partir do século XXI, é possível observamos mudanças gradativas em torno da IA, ocorrento, conseqüentemente, uma curiosidade em torno do tema. Ahmand *et al.*, (2021) em seu estudo, observaram que o crescimento científico foi impulsionado pelo aumento nos investimentos em pesquisa, reforma educacional e desenvolvimento de novos algoritmos para popularização da IA.

Dave e Patel (2023) abordam que o uso de algoritmos de IA no processo de publicação pode torná-lo mais eficiente, reduzindo a quantidade de trabalho dos revisores no processo duplo cego, o que pode permitir tempos de publicação mais céleres e com uma maior eficiência. Vale ressaltar que a integração tecnológica com a IA pode ser um fator para explicar o aumento de publicações científicas e conseqüentemente um aumento no número de pesquisas sobre IA (Pan, 2016).

**Gráfico 1:** Distribuição do número de artigos recuperados na Science Direct sobre IA e ética no período de 2019 a 2023.



Fonte: Elaborado pelos autores (2023).

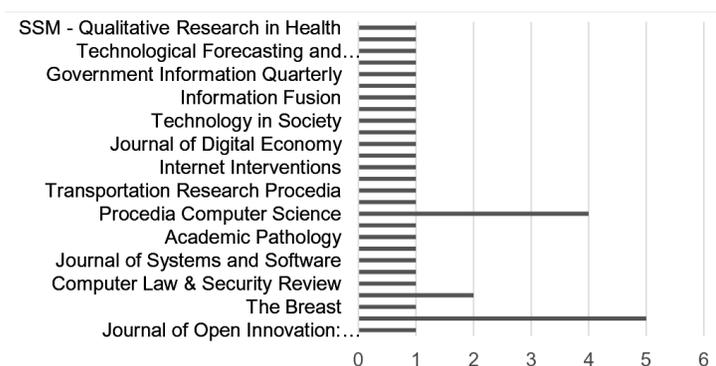
As discussões direcionadas à inteligência artificial, seja na área da saúde, tecnologia ou educação, têm demonstrado tanto entusiasmo sobre os potenciais benefícios da IA como preocupação sobre potenciais riscos (Carter, 2020). Russell, Dewey e Tegmark (2015) apontam que a busca pela IA tem o potencial de trazer benefícios sem precedentes para a humanidade e um dos grandes desafios éticos na atualidade está relacionado ao fato dos algoritmos da IA serem utilizados para atividades com dimensões sociais e cognitivas, antes realizadas por humanos.

Conforme Cuellar (2023), a capacidade da IA de realizar tarefas antes executadas por seres humanos reverbera no meio social, acadêmico e científico, assim as discussões

sobre os dilemas éticos têm sido uma constante. É importante e necessário salientar que estudos sejam pautados nos diversos âmbitos, e estes reportem os impactos dos sistemas da IA no cotidiano de indivíduos, organizações e governos. Atualmente, as preocupações com a IA variam desde a proteção de dados, segurança da informação, até na tomada de decisão na área médica (González Esteban; Calvo, 2022).

A falta de padrões objetivos universais para tomada de decisões pode inicialmente parecer pouco adequada à IA, portanto, é de suma importância regulamentar a utilização da IA para mitigar os riscos e os impactos negativos, incluindo o uso indevido, à medida que a IA evolui, trazendo assim benefícios para sociedade (Macintyre *et al.*, 2003).

**Gráfico 2:** Número de publicações distribuída por periódicos internacionais



Fonte: Elaborado pelos Autores (2023).

Os 03 (três) periódicos que apresentaram um número maior de publicações sobre o tema IA e ética, foram (i) *The ORBIT Journal*, (ii) *Technology in Society* e o *Procedia Computer Science* com 5, 2 e 4 artigos publicados, respectivamente. Entretanto, em termos do fator centimétrico de citação, no *Journal Impactor Factor (JIF)* que reflete o número médio de citações que os artigos receberam nesse periódico, o *Procedia Computer Science* apresenta um JIF de 2,56 enquanto o *The ORBIT Journal* e o *Technology in Society* apresentam um JIF de 0,95 e 1,81, respectivamente. As publicações sobre a temática IA e ética, vem sendo discutidas no meio científico, as preocupações éticas e os direitos dos seres humanos em detrimento das IA é tópico de debate público, governamental e de instituições responsáveis pela implantação e implementação dos regulamentos que norteiam a prática e o uso das IA (Miller; Ohrvik; Coldicutt, 2018).

## 4 Considerações finais

Conclui-se que as considerações éticas na área tecnológica são pouco pesquisadas e incoerentes, o que requer maiores investigações à medida que a tecnologia avança.

A IA por ser um campo de estudo em desenvolvimento, as tomadas de decisões visam abordar as implicações morais, sociais e legais do desenvolvimento sobre a IA. Desse modo, sugere-se que pesquisas futuras possam afunilar a discussão sobre a inteligência artificial e, para isso, pode-se realizar uma revisão de escopo, a fim de responder às lacunas científicas de uma determinada área conhecimento, abordando conceitos e definições teóricas.

## Contribuição dos autores

**Delineamento dos estudos de base, a realização da pesquisa e a redação do manuscrito:** Sheilla da Silva Barroso.

**Delineamento da pesquisa e a revisão crítica do manuscrito:** Laurides Batista Cruz e Daniele Carvalho Castro.

## Referências

AHMAD, Shakil; UR REHMAN, Shafiq; ASHIQ, Murtaza. A bibliometric review of Arab world research from 1980-2020. **Science & Technology Libraries**, [s. l.], v. 40, n. 2, p. 133-153, 2021.

ARAÚJO, Carlos A. A. Bibliometria: evolução histórica e questões atuais. **Em questão**, [s. l.], v. 12, n. 1, p. 11-32, 2006.

CARTER, Stacy M. *et al.* The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. **The Breast**, [s. l.], v. 49, p. 25-32, 2020.

CUELLAR, Michael. A virtue ethical approach to the use of artificial intelligence. **Data and Information Management**, [s. l.], p. 100037, 2023.

DAVE, Manas; PATEL, Neil. Artificial intelligence in healthcare and education. **British Dental Journal**, [s. l.], v. 234, n. 10, p. 761-764, 2023.

DÍAZ-RODRÍGUEZ, Natalia *et al.* Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. **Information Fusion**, [s. l.], v.99, p. 1-24, 2023.

EVGRAFOVA, Irina *et al.* Ethical problems of practical interaction between strong artificial intelligence and humans in the urban transport environment and legal proceedings. **Transportation Research Procedia**, [s. l.], v. 63, p. 2094-2098, 2022.

DE FELICE, Fabio *et al.* Artificial Intelligence or Augmented Intelligence? Impact on our lives, rights and ethics. **Procedia Computer Science**, [s. l.], v. 200, p. 1846-1856, 2022.

FIGUEREDO, Wilton Nascimento *et al.* Análise bibliométrica da produção brasileira sobre a Covid-19. **Revista Baiana de Enfermagem**, Salvador, v. 34, p.1-10, 2020.

GONZÁLEZ ESTEBAN, Elsa; CALVO, Patrici. Ethically governing artificial intelligence in the field of scientific research and innovation. **Heliyon**, [s. l.], v. 8, n. 2, p.1-9, 2022.

MACINTYRE, Michael R. *et al.* Ethical considerations for the use of artificial intelligence in medical decision-making capacity assessments. **Psychiatry research**, [s. l.], v. 328, p. 1-6, 2023.

MILLER, Catherine; OHRVIK-STOTT, Jacob; COLDICUTT, Rachel. **Regulating for responsible technology: capacity, evidence and redress: a new system for a fairer future.**, London: Doteveryone, 2018.

PAN, Yunhe. Heading toward artificial intelligence 2.0. **Engineering**, [s. l.], v. 2, n. 4, p. 409-413, 2016.

RUSSELL, Stuart; DEWEY, Daniel; TEGMARK, Max Erik. Research priorities for robust and beneficial artificial intelligence. **AI Magazine**, [s. l.], v. 36, n. 4, p. 105-114, 2015.

XUE, Lan; PANG, Zhenjing. Ethical governance of artificial intelligence: An integrated analytical framework. **Journal of Digital Economy**, [s. l.], v. 1, n. 1, p. 44-52, 2022.

WANG, Hanchen *et al.* Scientific discovery in the age of artificial intelligence. **Nature**, [s. l.], v. 620, n. 7972, p. 47-60, 2023.