

# Unveiling Ethical AI: An In-Depth Bibliometric and Visual Exploration

Qihui Yuan<sup>1,2</sup>, Tingxuan Chen<sup>3</sup>, Qing Yang<sup>2</sup>, Zongming Zhang<sup>3,\*</sup>

<sup>1</sup>Institute of Literature in Chinese Medicine, Nanjing University of Chinese Medicine, Nanjing, CHINA.

<sup>2</sup>College of Traditional Chinese Medicine, Ningxia Medical University, Yinchuan, CHINA.

<sup>3</sup>School of Marxism and Medical Humanities, Nanjing University of Chinese Medicine, Nanjing, CHINA.

## ABSTRACT

This study meticulously delineates the research progression, focal points, developmental path and emergent evolutionary trajectories of computational intelligence within the domain of ethics. A comprehensive quantitative report, embodying bibliometric principles, was prepared using the Web of Science Core Collection database. The methodology involved conducting an extensive search for original articles and reviews related to "ethical artificial intelligence". Relevant data from these sources were extracted and integrated into CiteSpace and Vosviewer for bibliometric and knowledge mapping analysis. The R package was used as an auxiliary tool in this process, facilitating the creation of co-occurrence networks and providing necessary information on authors, countries/regions, institutions, documents, references and keywords. This investigation includes an aggregate of 327 articles. The research on ethical AI originated in 2001 and experienced a marked escalation from 2015 onwards. Luciano Floridi has made the most significant contribution with regard to the volume of scholarly outputs. The USA emerged as the region with the most prolific output and the University of Oxford stands as the institution contributing the greatest amount of articles. The periodical registering the most prolific scholarly contributions is "Science and Engineering Ethics". Keyword analysis reveals that, in addition to topic-specific keywords such as "artificial intelligence", "big data" and "artificial neural network", there are also keywords that align with ethical theory research, such as "ethic issue" and "medical ethics", as well as those congruent with technical application research, including "technology", "model" and "design". These keywords substantially shape the principal directions of future research in this domain. The study of ethics in the field of AI is undergoing significant growth and holds immense potential. It is imperative to foster international and interdisciplinary collaboration among nations and institutions in order to advance further. As AI theories and technologies continue to mature, research addressing the ethical dimensions of AI, its technical applications, or the integration of both, will undoubtedly become key areas of sustained development and innovation in the future.

**Keywords:** Bibliometric analysis, Artificial intelligence, Ethical issue, Citespace, VOSviewer.

## Correspondence:

### Zongming Zhang

MD, School of Marxism and Medical Humanities, Nanjing University of Chinese Medicine, 138 Xianlin Avenue, Qixia District, Nanjing-210023, CHINA.  
Email: 13952092536@163.com

**Received:** 13-04-2024;

**Revised:** 16-05-2024;

**Accepted:** 06-06-2024.

## INTRODUCTION

Artificial Intelligence (AI) has become ubiquitous in our society, impacting various aspects of our daily lives. From complex surgical procedures to lesion detection and even smartphones, AI encompasses a range of technologies including intricate computational formulas, machine-guided learning processes, profound neural architectures and knowledge transference methodologies, to name a few constituent elements.<sup>1</sup> While AI finds applications in diverse fields, its core essence still lies in algorithms and machine learning. However, what distinguishes

artificial intelligence from other sophisticated technological innovations is its unparalleled capacity to augment its foundational coding and decision-strategizing mechanisms via profound learning competencies. This distinction underscores the ethical considerations associated with AI, with discretion, safeguarding, duty, equitability, answerability and lucidity emerging as prominent concerns.<sup>2</sup>

The emergence of ethical issues relevant to the utility of AI in various domains has been increasing continuously. In order to enhance the quality of healthcare services and improve the efficiency of healthcare resources, AI technology has been widely adopted in healthcare institutions. AI-supported technologies rely on extensive medical research and patient treatment records for learning, playing a significant role in strengthening physician diagnosis and treatment decision-making processes. However, the integration of artificial intelligence within healthcare provision



DOI: 10.5530/ijper.58.4.120

### Copyright Information :

Copyright Author (s) 2024 Distributed under  
Creative Commons CC-BY 4.0

**Publishing Partner :** EManuscript Tech. [www.emanuscript.in]

elicits cross-border considerations of an ethical, juridical, societal and commercial nature.<sup>3</sup> Notable concerns in this context include privacy, network security, data integrity, data ownership, medical ethics, responsibility for medical errors and the risk of system failures. Given the nature of healthcare service systems, ethical issues in medicine have become the most pressing challenge. AI technology has the potential to compromise patient preferences, safety and privacy. However, current policies and ethical guidelines governing healthcare services incorporating AI and its applications lag behind the pace of AI's development. Artificial intelligence introduces ethical quandaries that hinder the advancement of its implementation in the domain of healthcare.<sup>4</sup> In the ever-evolving contemporary business environment characterized by globalization, openness and competitiveness, the field of management accounting has emerged as a successful example of artificial intelligence application in operational and decision-making processes. The advancements brought by AI solutions have provided organizations with significant advantages across various domains. However, the use of AI capabilities also raises certain professional ethical concerns. The implementation of AI solutions in organizational activities and processes gives rise to ethical issues related to the security, confidentiality, transparency and integrity of AI solutions.<sup>5</sup> Therefore, neglecting the ethical issues associated with AI solutions can lead to collective problems.<sup>6</sup> The metamorphosis instigated by AI in the labor sphere has extensive societal repercussions. Systems augmented by artificial intelligence dictate the decisions pertaining to an individual's employment, advancement, or loan endorsement, in addition to curating the advertisements and journalistic pieces presented to consumers.<sup>7</sup> Such algorithmic determinations may result in inequitable adverse outcomes or potentially infringe upon fundamental human rights.<sup>8</sup> The pace of technological advancement and our lifestyles hinge on this comfort zone, facilitated by AI's capacity to autonomously oversee numerous selections and decision-making procedures, culminating in a semblance of personal agency. This reliance and personal state elevate AI to a position of heightened influence and potency. Consequently, it becomes imperative to establish requisite ethical guidelines.<sup>2</sup>

Bibliometric analysis, serving as a method that allows for the quantitative and qualitative analysis and visualization of all publications in a given research field, has found widespread application in the medical domain.<sup>9</sup> Yet, up until the present moment, there is an evident lack of bibliometric analysis associated with ethical AI. This study aims to achieve a comprehensive understanding of the progress and various foundational aspects of research in the field of ethics as it pertains to artificial intelligence. Utilizing the Web of Science Core Collection database, it performs a bibliometric analysis on the relevant articles. The study encompasses a thorough examination of the related authors, nations/regions, institutions, journals, keywords, references and co-occurrence networks. Through a systematic examination of

these aspects, this investigation furnishes a holistic examination of the evolving landscape of ethical AI, identifying key areas of activity, potential future development trajectories and potential emerging issues.

## MATERIALS AND METHODS

### Inclusion and Exclusion Criteria

#### *Inclusion Criteria*

(1) Literature related to the research on artificial intelligence in the field of oral medicine; (2) The type of the literature being "Article" and "Review"; (3) The language of the literature being English.

#### *Exclusion Criteria*

Literature with incomplete or missing information.

### Literature Retrieval Strategy

The Web of Science (WoS) Core Collection database is leveraged as the foundational basis for this investigation. The retrieval strategy combines keywords related to artificial intelligence with those relevant to the field of oral medicine. After reviewing the literature and performing a preliminary search, the defined retrieval strategy is: TS=(“artificial intelligence” OR AI OR “artificial intelligent” OR “artificial intelligence technology” OR “computational intelligence” OR “machine intelligence” OR “computer reasoning” OR “computer vision systems” OR “computer vision system” OR “deep learning” OR “deep network” OR “neural network”) AND TS=(“ethical issues” OR “ethical issue” OR “moral policy” OR “moral policies” OR “medical ethics” OR “clinical ethics” OR “professional ethics”) The literature timeline is set from the inception of the database to December 31st, 2023. Following the inclusion and exclusion criteria, the literature is screened and the full records and cited references of the target literature are extracted (Figure 1). They are saved locally as plain text files, following the naming convention "download\_\*.txt".

### Methods

R software (version 4.1.1, Auckland University, New Zealand, <https://www.r-project.org/>) (bibliometrix package), Gephi Software (version 0.9.5, Compiègne, France, <https://gephi.org>), VOSviewer (version 1.6.18, Leiden University, the Netherlands, <https://www.vosviewer.com/>) and CiteSpace (version 5.8.R3, Drexel University, Philadelphia, PA, USA, <https://citespace.podia.com/>) were employed for data manipulation and graphical representation.<sup>10-12</sup>

## RESULTS

### Interpretation of Maps and Indicators

Within the CiteSpace network maps, each node represents different types of research (co-authorship/co-citation/

co-occurrence). The diameter of the node signifies the frequency of publications; the different color rings on the nodes, progressing from inner to outer, symbolize the chronology of publications from past to present. The color starting at the ring indicates the publication situation since that time and the width of the ring denotes the volume of publications within that temporal node. The chromatic attributes of the connections among nodes denote the annum of the inaugural partnership between the investigative entities.

Betweenness centrality, a notion postulated by the American sociologist, Linton Freeman, alludes to the frequency with which a node serves as an intermediary conduit on the most direct trajectory between two distinct nodes. A higher frequency of acting as a bridge equates to greater betweenness centrality, implying that the node has more influence within the collaborative network. In CiteSpace, nodes with a betweenness centrality exceeding 0.1 are referred to as key nodes and are depicted with a purple outer circle on the map.<sup>13</sup> Modularity measures network modularization, with higher values indicating better network clustering. Q values range from 0 to 1; when  $Q > 0.3$ , it suggests a significant community structure within the network. The Silhouette value, proposed by Kaufman and Rousseeuw in 1990, assesses clustering effects by evaluating clustering efficiency and network uniformity. A silhouette value nearing 1 signals pronounced network homogeneity, a silhouette coefficient of 0.7 signifies a clustering outcome with substantial certainty and values exceeding 0.5 suggest a clustering verdict that is deemed acceptable.<sup>12,14</sup>

In the VOSviewer network depiction, entities are symbolized by their identifiers, which conventionally are illustrated as circles. An entity's circle and identifier's magnitude is influenced by its inherent weight. A pronounced weight for an entity manifests in a more expansive circle and identifier. Occasionally, identifiers might be omitted to counteract any superimposition of labels. An entity's hue is designated by the cluster it associates with. Connective lines between entities symbolize relationships.

Visually, the spatial separation between two nodes roughly conveys the affinity of journals based on co-citation connections. Typically, the more proximate the nodes, the more potent their mutual relevance. Entities are symbolized by their labels in a method paralleling that of network and overlay depictions. Each locus in the item density portrayal is colored, indicating item concentration at that locus. Conventionally, the hue progression is from blue, through green, to yellow. The denser the item distribution around a particular locus, especially with greater item weights, the more the hue skews towards yellow. The color intensity for a particular cluster is ascertained by the prevalence of entities from that cluster proximate to the locus. Analogous to item density portrayal, the significance of each entity also factors into the analysis.<sup>15</sup>

### Worldwide patterns of scholarly outputs and citation frequencies

Figure 1 reflects the annual scientific research activity and stages of scientific research development in the field of AI focused on medical ethics. After data retrieval and screening, a total of 327 papers were collected, including 241 original articles and 86 reviews. The study of AI medical ethics began in 2001 and since then, the annual publication volume and the annual cumulative publication volume have been steadily increasing at an annual growth rate of 6.25%. According to this development trend, the entire development stage can be initially divided into two stages. The first stage (initial development stage) spans from 2001 to 2015, characterized by: A very small annual publication volume and a very slow rise in the curve of the annual cumulative publication volume; Incomplete annual publications during this stage, with no relevant research published in seven years; A single form of publication, generally original articles. The second stage (steady development stage) spans from 2017 to May 2023, characterized by: A significant increase in the annual publication volume and a faster rise in the curve of the annual cumulative publication volume; Complete annual publications during this stage, with no gap years appearing; Diversified forms of publication. In

**Table 1: Top 10 most productive countries/regions.**

| Rank | Country/Region  | Count | Total Citation | Average Article Citations | Centrality | h-index |
|------|-----------------|-------|----------------|---------------------------|------------|---------|
| 1    | USA             | 89    | 1057           | 16                        | 0.25       | 22      |
| 2    | England         | 57    | 964            | 28.4                      | 0.23       | 16      |
| 3    | Germany         | 38    | 383            | 18.2                      | 0.17       | 13      |
| 4    | China           | 32    | 127            | 4.9                       | 0.11       | 7       |
| 5    | Australia       | 30    | 387            | 20.4                      | 0.08       | 10      |
| 6    | Italy           | 30    | 351            | 20.6                      | 0.08       | 11      |
| 7    | Canada          | 30    | 198            | 19.8                      | 0.12       | 11      |
| 8    | The Netherlands | 23    | 173            | 17.3                      | 0.06       | 10      |
| 9    | France          | 22    | 200            | 15.4                      | 0.1        | 9       |
| 10   | Switzerland     | 19    | 67             | 7.4                       | 0.03       | 7       |

addition to original articles, the number of review articles also shows a trend of substantial annual growth (Figure 2A). Figure 2B shows the trend of quantity changes in AI-related studies in general medical fields and professional medical ethics research. The development trends for both areas are generally aligned, with each experiencing a period of stagnation followed by rapid

development. This pattern reflects a synergy and consistency between the two fields. Specifically, the quantity of AI research papers in the broader medical field has experienced substantial growth since 2017, culminating in a peak in 2022. Concurrently, research on AI in the realm of professional medical ethics has also been advancing and progressing in parallel.

**Table 2: Top 10 most productive institutions.**

| Rank | Institution              | Count | Centrality | Total Link Strength | Local Citation | Total Citation | Average Article Citation | <i>h</i> -index |
|------|--------------------------|-------|------------|---------------------|----------------|----------------|--------------------------|-----------------|
| 1    | Univ Oxford              | 15    | 0.04       | 14                  | 521            | 560            | 37.33                    | 8               |
| 2    | Univ Melbourne           | 6     | 0          | 6                   | 229            | 232            | 33.14                    | 6               |
| 3    | Harvard Med Sch          | 6     | 0.14       | 7                   | 122            | 122            | 17.43                    | 4               |
| 4    | Univ Toronto             | 6     | 0.1        | 6                   | 150            | 150            | 21.43                    | 5               |
| 5    | Zhejiang Univ            | 6     | 0.05       | 8                   | 138            | 139            | 19.86                    | 5               |
| 6    | Delft Univ Technol       | 6     | 0.01       | 3                   | 99             | 103            | 17.17                    | 4               |
| 7    | Univ Leeds               | 6     | 0.09       | 6                   | 15             | 15             | 2.5                      | 3               |
| 8    | Univ Basel               | 5     | 0          | 8                   | 90             | 90             | 15                       | 4               |
| 9    | Univ Utrecht             | 5     | 0          | 3                   | 177            | 181            | 30.17                    | 5               |
| 10   | Univ Calif San Francisco | 5     | 0.14       | 4                   | 80             | 80             | 16                       | 4               |

**Table 3: The ten foremost prolific writers and the uppermost ten co-referenced authors amassing the most substantial citation counts.**

| Rank | Authors              | Count | Total Citations | <i>h</i> -index | <i>g</i> -index | Cited-authors       | Count | Total Citations | Total Link Strength | Centrality |
|------|----------------------|-------|-----------------|-----------------|-----------------|---------------------|-------|-----------------|---------------------|------------|
| 1    | Floridi, Luciano     | 4     | 180             | 4               | 4               | FLORIDI L           | 64    | 93              | 1111                | 0.18       |
| 2    | Buyx, Alena          | 3     | 145             | 3               | 4               | MITTELSTADT BD      | 42    | 52              | 612                 | 0.03       |
| 3    | Li, Luming           | 2     | 7               | 1               | 2               | JOBIN A             | 38    | 39              | 376                 | 0.04       |
| 4    | Lorenzini, Giorgia   | 2     | 1               | 1               | 1               | ESTEVA A            | 27    | 30              | 231                 | 0.04       |
| 5    | Arabi, Hossein       | 2     | 7               | 1               | 2               | OBERMEYER Z         | 25    | 31              | 311                 | 0.05       |
| 6    | Taddeo, Mariarosaria | 2     | 143             | 2               | 2               | CHAR DS             | 25    | 33              | 355                 | 0.17       |
| 7    | Salimi, Yazdan       | 2     | 7               | 1               | 2               | TOPOL EJ            | 24    | 26              | 237                 | 0.04       |
| 8    | Brooks, Laurence     | 2     | 13              | 2               | 2               | MORLEY J            | 24    | 38              | 502                 | 0          |
| 9    | Shih, Po-Kang        | 2     | 13              | 1               | 2               | HAGENDORFF T        | 23    | 25              | 277                 | 0.08       |
| 10   | Ferdowsi, Sohrab     | 2     | 7               | 1               | 2               | EUROPEAN COMMISSION | 21    | 30              | 232                 | 0.01       |

## Analysis of co-authorship

### Authors

A total of 142 authors participated in AI research in the field of medical ethics. The visualization of the author collaboration network (Figure 3A, 3B) shows that the earliest authors to delve into this research area were Borenstein Jason<sup>16</sup> and Buyx Alena.<sup>17</sup> There are four small groups and one large group in the author collaboration network. The main and largest collaboration network is composed of people like Zaidi Habib, Shiri Isaac and Gunduz Deniz. Floridi Luciano is a connector in one of the small collaboration groups and also the author with the highest publication volume (Table 3). He co-authored a highly-cited paper in 2020 with authors like Taddeo Mariarosaria in his group.<sup>18</sup> This extensively referenced manuscript has garnered significant recognition within the domain of AI medical ethics studies and has accrued in excess of 96 citations thus far. Notably, although Buyx Alena is a prolific author with a substantial

volume of publications, her collaboration network is limited, showing minimal cooperation and communication with other authors and research groups. In general, the author collaboration network groups are relatively mature and dispersed, but there is still a lack of collaboration between groups. Among the top 10 authors in terms of publication volume, Floridi Luciano, Buyx Alena and Taddeo Mariarosaria are ranked in the top three, with a publication volume far exceeding other authors on the list and their h-index and g-index also decrease in order (Figure 3C). Both Floridi Luciano and Taddeo Mariarosaria are from the University of Oxford in the UK, while Buyx Alena is from Tech Univ Munich in Germany and these are also the institutions and nationalities of the top three authors in terms of publication volume in the field of human collective AI medical ethics. The main research areas of the three professors are digital data ethics, biomedical ethics and ethics of digital technologies, respectively. Despite not having a high publication volume, Laurence Brooks still ranks high in the h-index and g-index, indicating that while

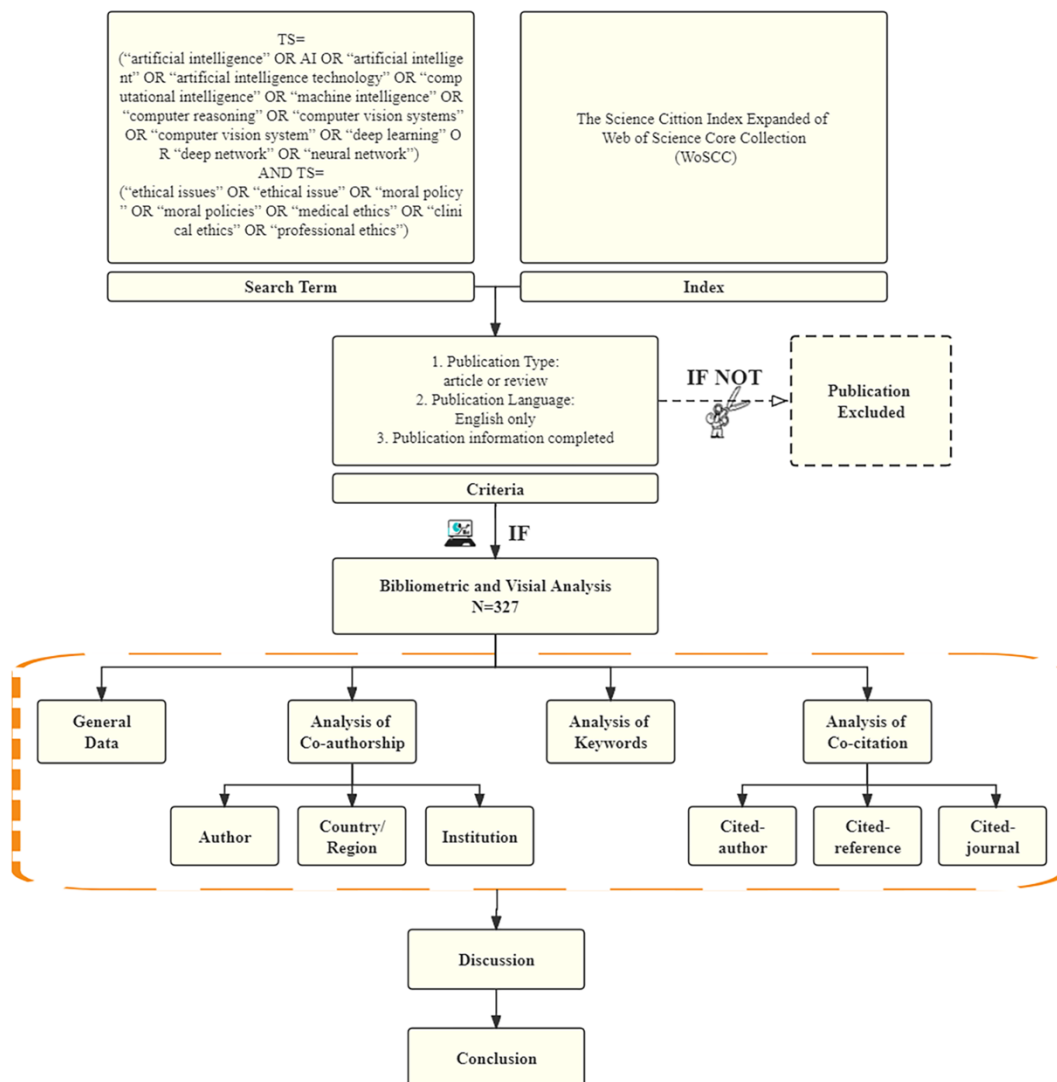


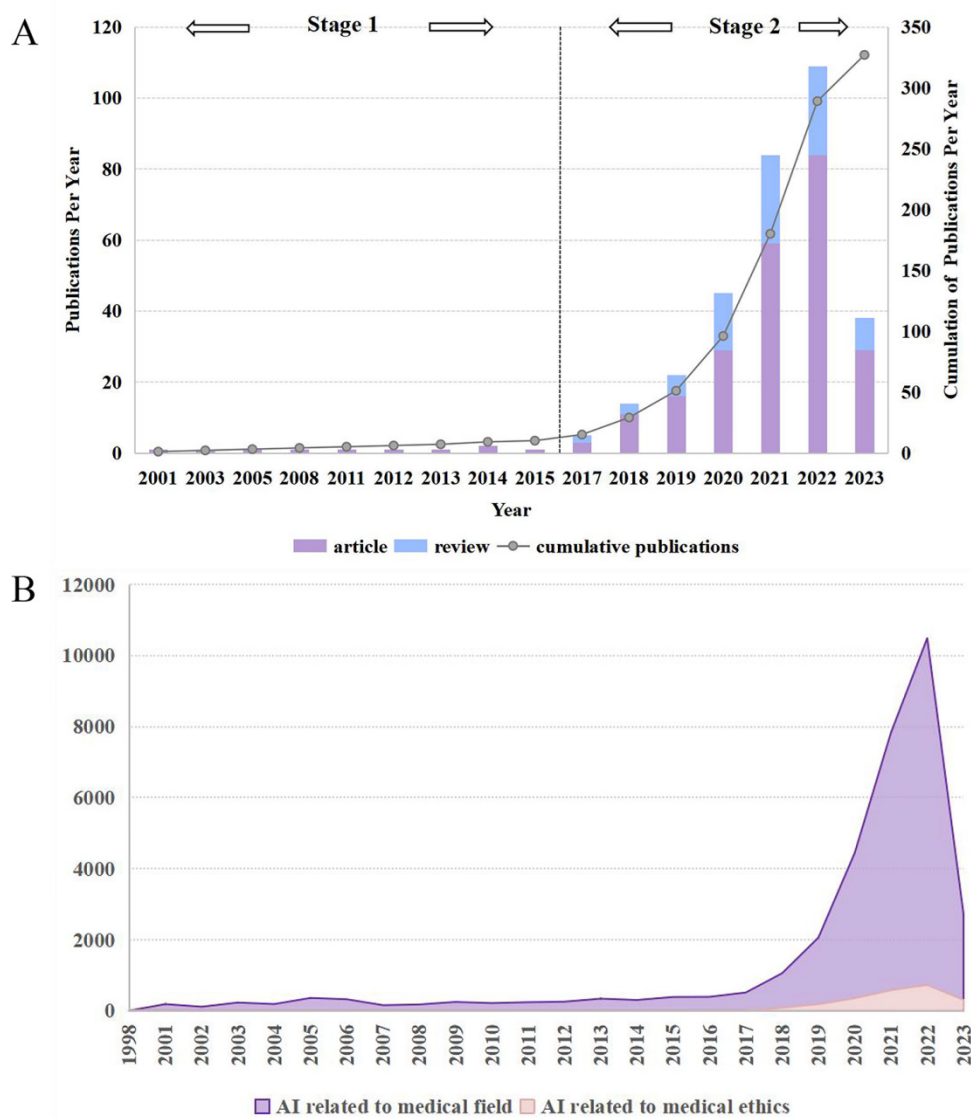
Figure 1: Research protocol.

the number of his articles is not high, their quality is excellent. *H-index*<sup>19</sup> is a composite index that can be used as an important indicator to evaluate the quantity and level of academic output of a researcher, a country, a journal, or an institution.<sup>20</sup> The *g-index* is a bibliometric measure that incorporates both the quality and quantity of a researcher's output, taking into account the citation distribution across all publications, thus providing a comprehensive assessment of their impact.<sup>21</sup>

### Country/Region

A total of 60 countries/regions have contributed to the field of AI in medical ethics research. A world geographical map was created to display the volume of papers in these countries/regions based on a color gradient. It is clear that the high-yield areas for papers in this field are concentrated in North America, with relatively less research in other regions (Figure 4A). As shown in Figure

4B, the high-yield countries/regions have close cooperation with each other. The United States maintains the most prolific collaborative endeavors with various nations, followed closely by Germany. Other high-yield countries like China, the United Kingdom and Italy have fewer international collaborations. Cooperation between other countries, especially developing countries, is relatively less. Upon organizing the nationalities of the corresponding authors of the publications, it is found that whether it is domestic or multinational cooperation, the United States is still in the lead, almost twice that of China. Europe is basically on par (Figure 4C). Table 1 lists the top 10 high-yield countries/regions in this field. The United States ranks first in terms of publication volume (89), total citations (1057), centrality (0.25) and *h-index* (22), indicating that it is the most advanced country in this field in the world. The United Kingdom closely follows in second place, but its average citation per paper is the highest (28.4), far exceeding that of the United States (16).

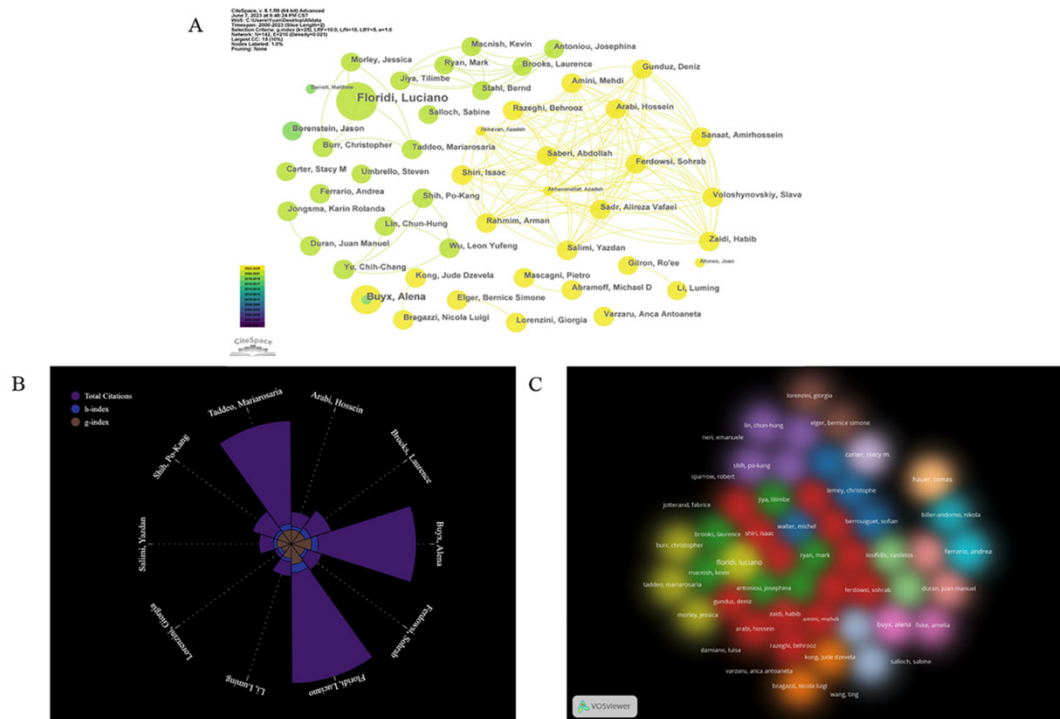


**Figure 2:** (A) Worldwide progression of publications focusing on AI-driven ethical inquiries from 2001 onwards. (B) The evolving composition of yearly publication volume commencing from 1998.

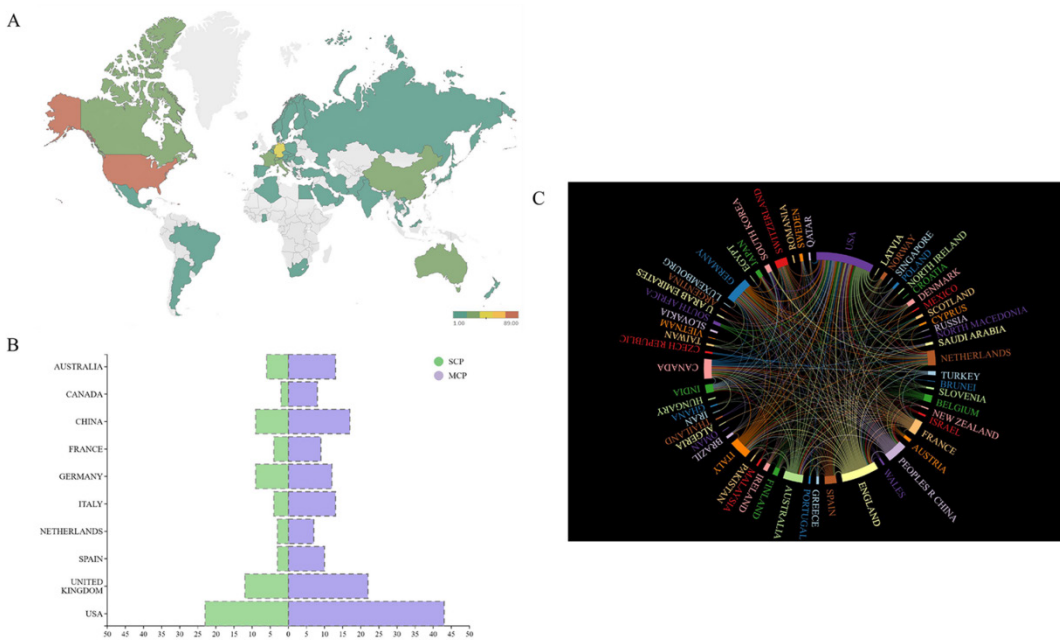
### Institution

A total of 200 institutions have conducted research in this field. In the top 10 institutions in terms of publication volume (Table 2), the University of Oxford ranks first in terms of publication volume (15), total link strength (14), local citation (521), total citation (560), average article citation (37.33) and h-index (8)

and its publication volume is 2-3 times that of other institutions on the list. The two research institutions with a centrality greater than 0.1 are Harvard Medical School (0.14) and the University of California, San Francisco (0.14), indicating that they have a great influence in the institutional cooperation network. This is also confirmed in the institution cooperation network diagram



**Figure 3:** (A) Co-authorship network of relevant authors. (B) Co-authorship cluster density of relevant authors. (C) Top 10 most productive authors from co-authorship network



**Figure 4:** (A) Cartographic representation delineating the aggregate scholarly outputs across distinct nations/regions. (B) Visual depiction of inter-country/region collaborations. The robustness of the connecting lines between nations signifies the intensity of their collaborative endeavors. (C) Most relevant countries/regions sorted by corresponding author.

(Figure 5A), which shows that the primary institutions involved are mainly based in the United States. These institutions are the most important and closely cooperating entities within the network. They also serve as key nodes in this network along with other institutions such as Imperial College London. Although the University of Oxford has the largest number of publications in recent years, its cooperation with other institutions is poor. Figure 5B shows the cooperation strength of each institution. Although the University of Oxford seldom cooperates, it has a strong cooperation with the Alan Turing Institute. Harvard Medical School has the strongest cooperation with Massachusetts Gene Hospital. Non-US institutions have relatively poor cooperation.

Funding institutions play a key role in conducting research and publishing articles. Through statistics, there are 3 funding institutions from the United States, 4 from the United Kingdom and one each from the European Commission, China and Australia. Among them, American institutions sponsor the most (29), followed closely by the United Kingdom (23) and then the European Commission (11), China (7) and Australia (5). This unequivocally suggests that the preeminent stance of the United States in this domain is intricately linked to its robust economic underpinning and backing (Figure 5C).

## Analysis of co-citation

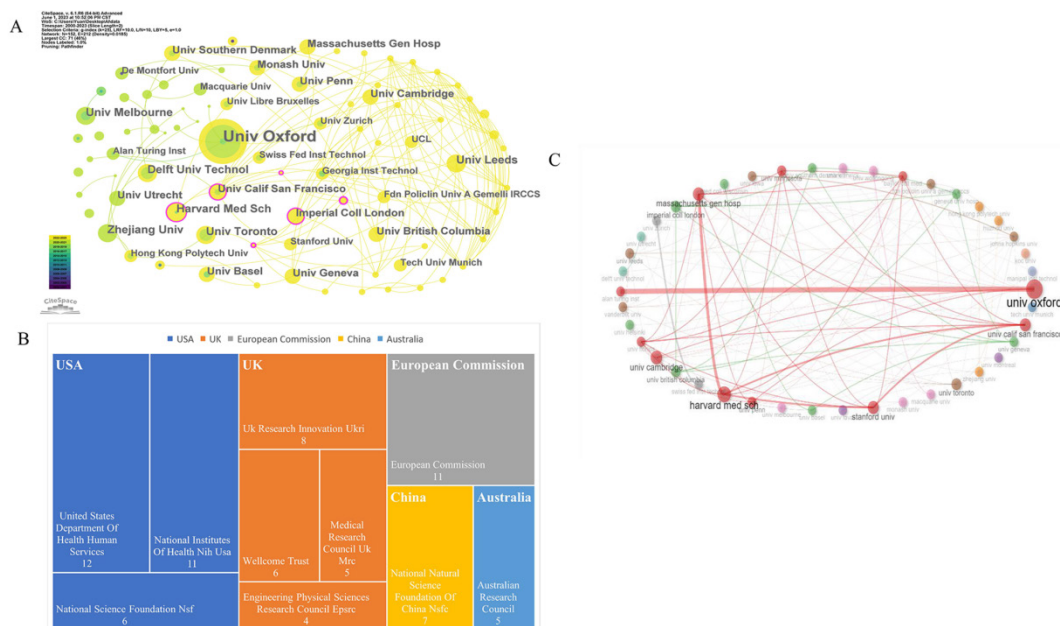
### Cited-author

A total of 456 co-cited authors have been included in the statistics. Table 3 not only displays the 10 most productive authors but also the top 10 co-cited authors with the highest citations. In the author co-citation network, the two largest groups are

represented by Floridi L and Esteva A, with total citations of 93 and 30 respectively (Figure 6A). Co-cited authors with a centrality greater than 0.1, besides Floridi L (0.18), include Char D (0.17), indicating that these two co-cited authors play a role as central nodes. In the entire author co-citation network, Floridi L, Jobin A and Hagendorff T, who are co-cited authors, are not only in the same group, but also have the closest relationship with each other, indicating that they are most commonly cited together in the same article (Figure 6B).

### Cited-journal

In this research, papers related to the study of medical ethics based on AI have been published in 219 academic journals, with these journals being co-cited at least 548 times. The recurrence of co-citations signifies the magnitude of a journal's imprint within a particular research domain, which in turn demarcates the journal's sway and prominence.<sup>22-24</sup> The journals with the most publications are SCIENCE AND ENGINEERING ETHICS (12), JOURNAL OF MEDICAL ETHICS (9) and WILEY INTERDISCIPLINARY REVIEWS-DATA MINING AND KNOWLEDGE DISCOVERY (9). The journals with the highest co-citation frequency are Science, Nature and Jama-j Am Med Assoc, with citation numbers of 105, 99 and 82 times respectively (Figure 7A, 7B). In the visualization map of Cited-journal in Figure C, it can be seen that: (1) High frequency co-cited journals have begun to appear in large numbers since 2015. (2) Red nodes represent journals that have been co-cited in large volumes in a short period of time. (3) Several journals have a centrality greater than 0.1, among which Science and Artificial Intelligence are the most representative, acting as key nodes. This suggests that the



**Figure 5:** (A) Co-authorship network of relevant institutions. (B) Collaboration Network of co-authorship of relevant institutions. The thickness of the line between countries reflects the frequency of the cooperation. (C) The most active funding agencies in relevant research.



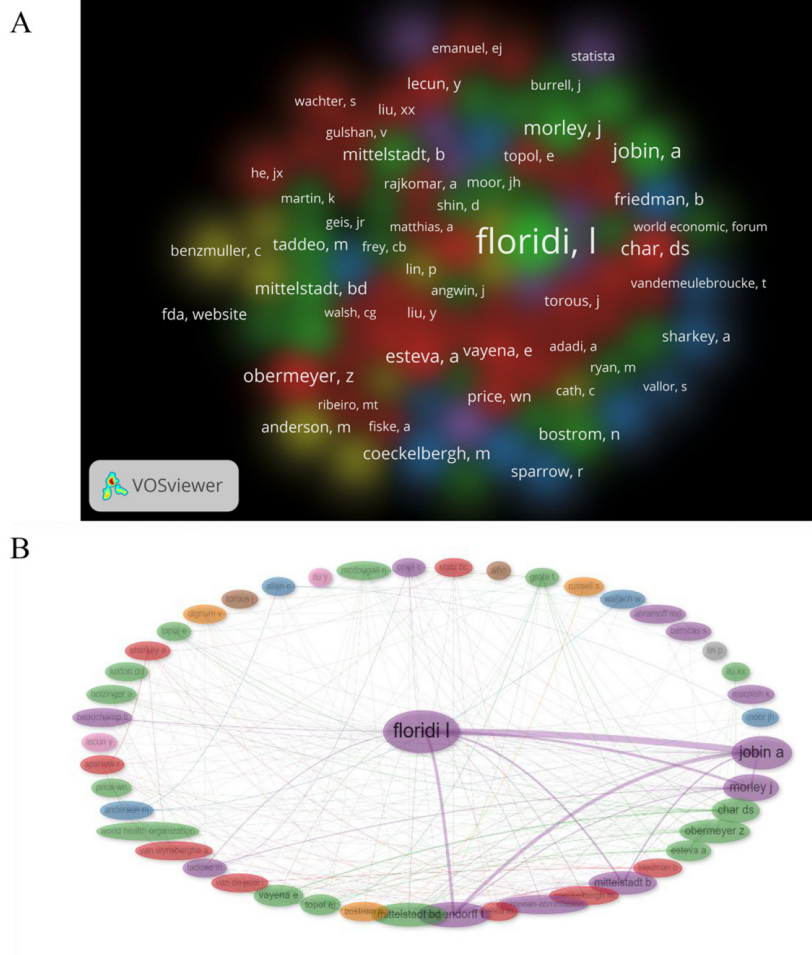
research field has received widespread attention in recent years and a large number of published papers have been cited, with the overall field starting to enter Stage 2, a steady development stage. To comprehensively decipher the knowledge interchange among varied disciplines and identify the vanguard or focal point of each field, the disciplinary dispersion of journals engaged in AI medical ethics studies was systematically cataloged. Using dual-map overlap (Figure 8), it was found that literature published in Medicine/Medical/Clinical or Psychology/Education/Health journals often cites Systems/Computing/Computer, Molecular/Biology/Genetics, Health/Nursing/Medicine, or Psychology/Education/Social. This suggests that the field is becoming more diversified with the overall development and increasingly has multidisciplinary characteristics.

## Reference

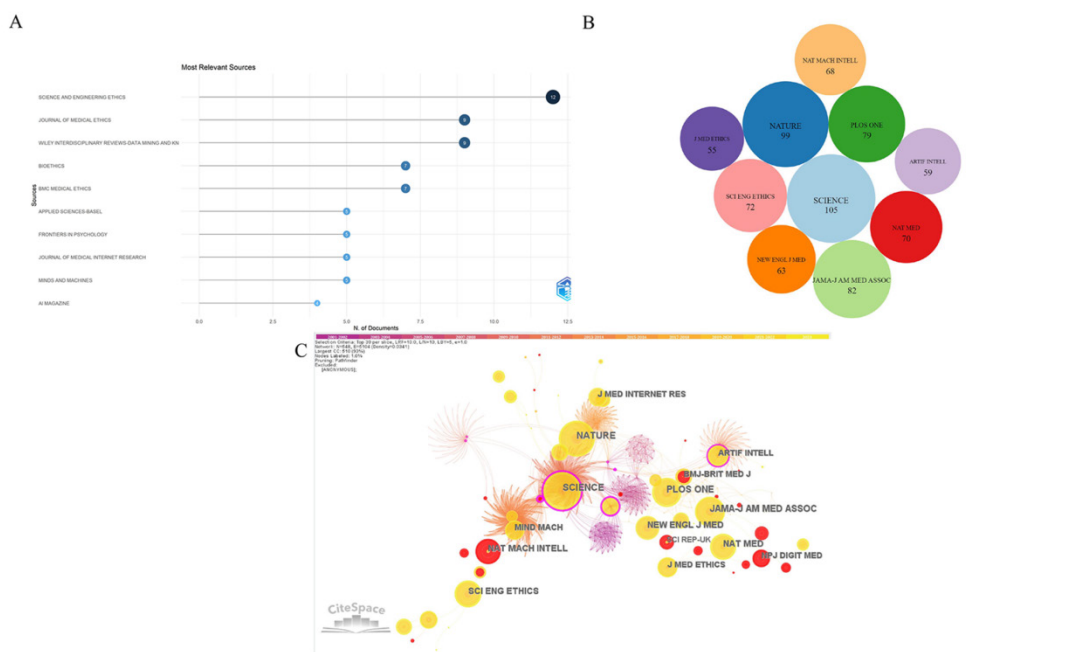
### Cited-reference

This study includes a total of 327 articles and 396 references. The articles with the most local citations are those published by MCDUGALL RJ and MITTELSTADT B in 2019 in the JOURNAL OF MEDICAL ETHICS and Nature Machine

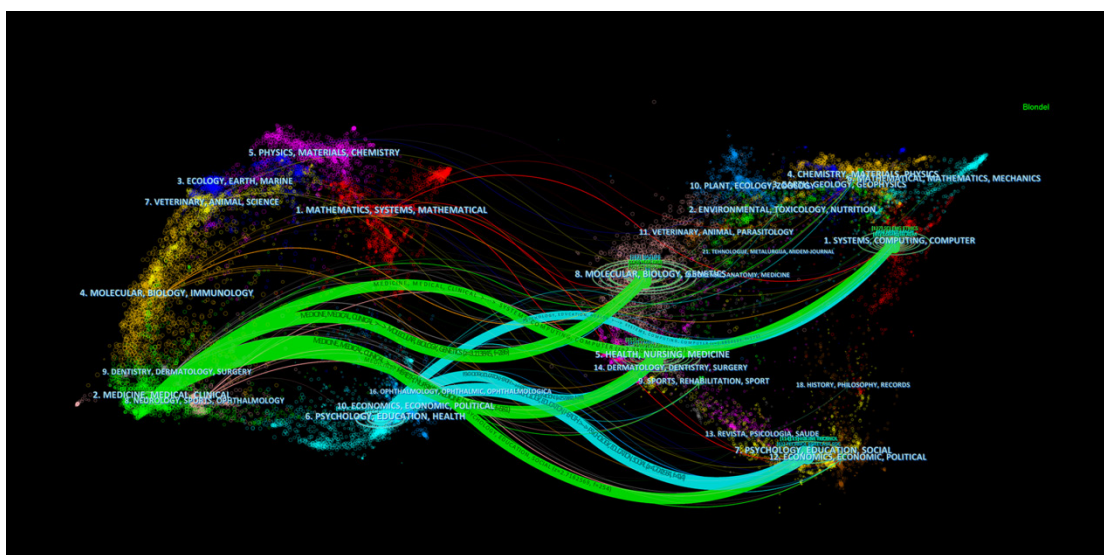
Intelligence, respectively, each cited 15 times (Figure 9A). The article with the most global citations is DWIVEDIYK's 2021 article titled "Setting the future of digital and social media marketing research: Perspectives and research propositions", published in the INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT. The article focuses on Information Science and Library Science and it has the highest total citations and total citations per year (334, 111.33) (Figure 9B). Table 4 shows the top 10 most locally cited references. The most co-cited reference is an original article "The global landscape of AI ethics guidelines", authored by Jobin A et al. and published in 2019 in the Nature Machine Intelligence journal. This article received 38 local citations and 679 global citations. Both Floridi *et al.*<sup>25</sup> in their 2018 article published in Minds and Machines and Topol *et al.*<sup>26</sup> in their 2019 article in Nature Medicine, had a centrality of 0.1. Combined with the Visualization map of Cited-reference (Figure 9C), the vast majority of local cited-reference nodes are red, indicating that these references were highly cited in a short period, most often after 2017. Figure 9D reveals the network map of co-cited references, showing that co-cited references mainly appear in four different group forms.



**Figure 6:** (A) Cluster density of cited-author. (B) Co-citation network of cited-author.



**Figure 7:** (A) Top 10 most relevant sources of AI ethics researches. (B) Top 10 cited-journals with most publication of AI ethics researches. (C) Visualization map of Cited-journal.



**Figure 8:** A dual-map overlap of journals on AI ethics researches.

### Historiography and Evolution

After applying a clustering algorithm, the timeline view of co-cited references intuitively reflects the temporal features of the research hotspots in this field (Figure 10A). The Modularity Q was 0.9406 and the mean Silhouette S was 0.9812, showing an excellent clustering effect and network homogeneity. Among the 10 clusters, "#5 Computational model" was the earliest research hotspot in this field, but there has been little research in this area since 2015. Subsequently, the new wave of "#0 Turing test" research has emerged and continues to this day, remaining at the center of research hotspots. At present, most of the key articles are still frequently cited, which suggests that research on

AI-driven medical ethics will persistently emerge as a focal point of investigation in the ensuing years.

According to historiography, Luxton DD's paper titled "Recommendations for the ethical use and design of artificial intelligent care providers"<sup>27</sup> published in ARTIFICIAL INTELLIGENCE IN MEDICINE in 2014, left a significant mark in the field of AI medical ethics, laying the groundwork for future research. Among the main research directions in this field, the paper titled "Computer knows best? The need for value-flexibility in medical AI"<sup>19</sup> by MCDUGALL RJ published in JOURNAL OF MEDICAL ETHICS in 2019 has become an essential part of recent research, as evidenced by both its publication network relationships and its local cited-documents (see Section 3.4.3.1

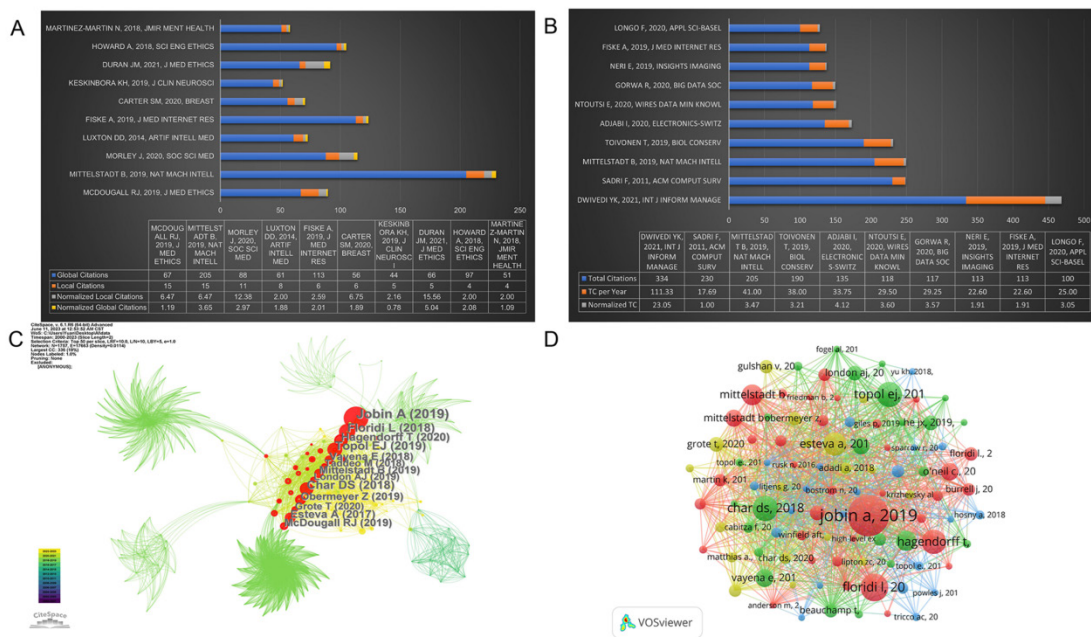


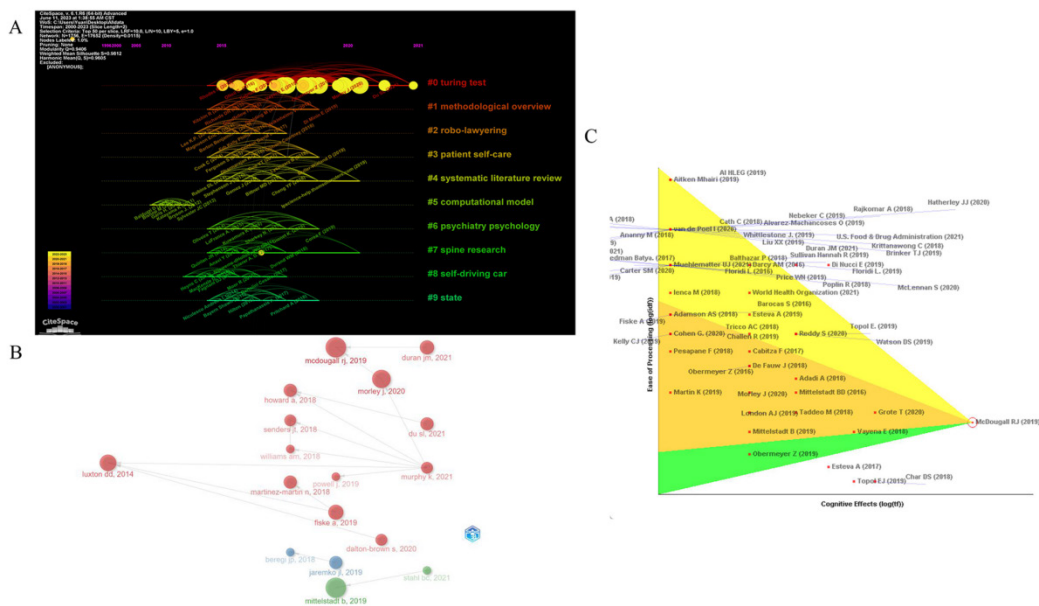
Figure 9: (A) Top10 local cited-documents. (B) Top10 global cited-documents. (C) Visualization map of Cited-reference. (D) Network map of Cited-reference.

Table 4: Top 10 most local cited-reference.

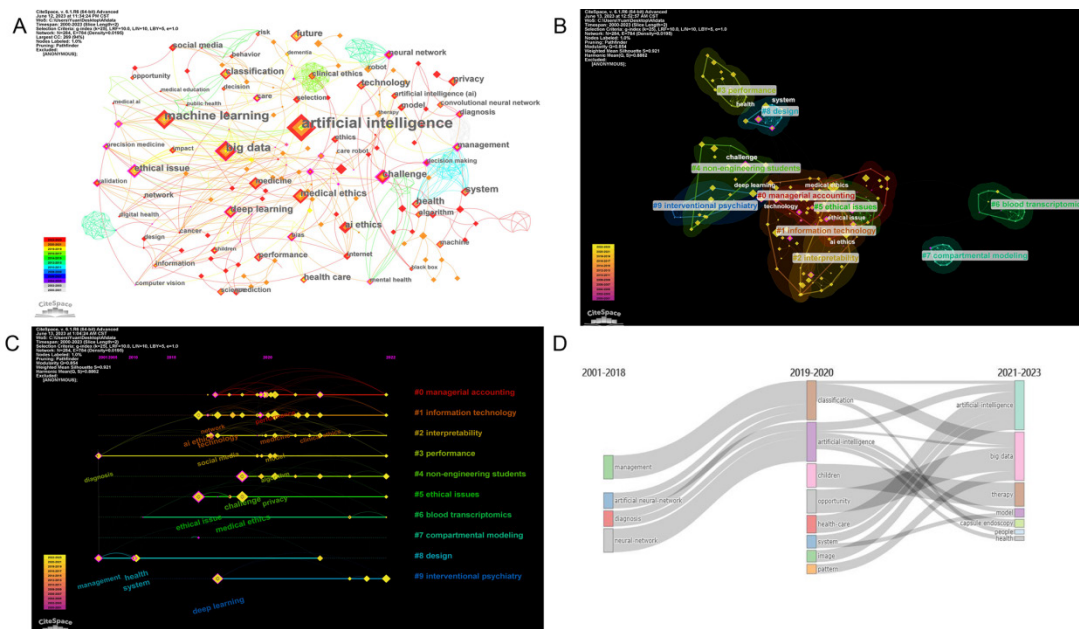
| Rank | Cited-reference   | Year | Citation | Centrality | Total Link Strength |
|------|---|------|----------|------------|---------------------|
| 1    | Jobin A, 2019, NAT MACH INTELL, V1, P389, DOI: 10.1038/s42256-019-0088-2        | 2019 | 38       | 0.09       | 202                 |
| 2    | Floridi L, 2018, MIND MACH, V28, P689, DOI: 10.1007/s11023-018-9482-5           | 2018 | 25       | 0.1        | 138                 |
| 3    | Char DS, 2018, NEW ENGL J MED, V378, P981, DOI: 10.1056/NEJMp1714229            | 2018 | 23       | 0.07       | 187                 |
| 4    | Topol EJ, 2019, NAT MED, V25, P44, DOI: 10.1038/s41591-018-0300-7               | 2019 | 23       | 0.1        | 172                 |
| 5    | Hagendorff T, 2020, MIND MACH, V30, P99, DOI: 10.1007/s11023-020-09517-8        | 2020 | 22       | 0.08       | 131                 |
| 6    | Esteva A, 2017, NATURE, V542, P115, DOI: 10.1038/nature21056                    | 2017 | 20       | 0.06       | 138                 |
| 7    | Obermeyer Z, 2019, SCIENCE, V366, P447, DOI: 10.1126/science.aax2342            | 2019 | 18       | 0.04       | 111                 |
| 8    | Vayena E, 2018, PLOS MED, V15, P0, DOI: 10.1371/journal.pmed.1002689            | 2018 | 16       | 0.03       | 121                 |
| 9    | McDougall RJ, 2019, J MED ETHICS, V45, P156, DOI: 10.1136/medethics-2018-105118 | 2019 | 15       | 0.04       | 125                 |
| 10   | Mittelstadt B, 2019, NAT MACH INTELL, V1, P501, DOI: 10.1038/s42256-019-0114-4  | 2019 | 15       | 0.01       | 105                 |

**Table 5: The 10 most influential papers in AI ethics historiography.**

| No. | Paper  | Title  | Some Author Keywords and Keywords Plus  | Year | LCS | GCS |
|-----|--|--|---|------|-----|-----|
| 1   | Luxton Dd, 2014, <i>Artif Intell Med</i><br>DOI:10.1016/J.Artmed.<br>2014.06.004   | Recommendations For The Ethical Use And Design Of Artificial Intelligent Care Providers.   | Artificial Intelligent Agents; Ethics; Practice Guidelines; Care Providers; Mental Health.                    | 2014 | 8   | 61  |
| 2   | Senders Jt, 2018, <i>Acta Neurochir</i><br>DOI:10.1007/S00701-017-3385-8           | An Introduction And Overview Of Machine Learning In Neurosurgical Care.  | Artificial Intelligence; Neurosurgery; A Brain-Tumor Segmentation; Pattern-Analysis.                          | 2018 | 3   | 76  |
| 3   | Williams Am, 2018, <i>Physiol Genomics</i> DOI:10.1152/Physiolgenomic.00119.2017   | Artificial Intelligence, Physiological Genomics and Precision Medicine.  | Artificial Intelligence; Functional Genomics; Physiological Precision Medicine.                               | 2018 | 2   | 48  |
| 4   | Howard A, 2018, <i>Sci Eng Ethics</i><br>DOI:10.1007/S11948-017-9975-2             | The Ugly Truth About Ourselves And Our Robot Creations: The Problem Of Bias And Social Inequity.   | Artificial Intelligence; Implicit Bias; Design Ethics; Professional Ethics; Robot Ethics; Gender.             | 2018 | 4   | 97  |
| 5   | Martinez-Martin N, 2018, <i>Jmir Ment Health</i> DOI:10.2196/Mental.9423           | Ethical Issues For Direct-To-Consumer Digital Psychotherapy Apps: Addressing Accountability, Data Protection, And Consent.                   | Ethical Issues Telemedicine; Mental-Health Intervention; Adolescent.  | 2018 | 4   | 51  |
| 6   | Mcdougall Rj, 2019, <i>J Med Ethics</i><br>DOI:10.1136/Medethics-2018-105118       | Computer Knows Best? The Need For Value-Flexibility In Medical AI.   | Shared Decision-Making; Health-Care; Encounter; Model.  | 2019 | 15  | 67  |
| 7   | Fiske A, 2019, <i>J Med Internet Res</i><br>DOI:10.2196/13216                      | Your Robot Therapist Will See You Now: Ethical Implications Of Embodied Artificial Intelligence In Psychiatry, Psychology and Psychotherapy. | Artificial Intelligence; Ethics; Psychiatry; Medicine; Socially Assistive Robotics; Autism Spectrum Disorder. | 2019 | 6   | 113 |
| 8   | Powell J, 2019, <i>J Med Internet Res</i><br>DOI:10.2196/16222                     | Trust Me, I'M A Chatbot: How Artificial Intelligence In Health Care Fails The Turing Test.   | Artificial Intelligence; Machine Learning; Medical Informatics; Digital Health; Ehealth; Chatbots.            | 2019 | 2   | 28  |
| 9   | Morley J, 2020, <i>Soc Sci Med</i><br>DOI:10.1016/J.Socscimed.2020.113172          | The Ethics Of Ai In Health Care: A Mapping Review.   | Artificial Intelligence; Ethics; Health Policies; Legal; Prediction; Diagnosis.                               | 2020 | 11  | 88  |
| 10  | Dalton-Brown S, 2020, <i>Camb Q Healthc Ethic</i><br>DOI:10.1017/S0963180119000847 | The Ethics Of Medical Ai And The Physician-Patient Relationship.   | Medical Ai; Gdpr; Algorithm Bias; Care Robots.  | 2020 | 3   | 17  |



**Figure 10:** (A) Visualization map of timeline view. (B) Visualization map of historiography. (C) Pennant diagram for MCDougall RJ's document published in 2019.



**Figure 11:** (A) Network map of keywords related to AI ethics researches. (B) Cluster map of keywords of AI ethics researches. (C) Graphical representation in a timeline perspective. The progression of time is demarcated by lines of varying hues, with nodes on these lines signifying the emergence of keyword clusters. (D) The thematic evolution map of AI ethics researches.

above) (Figure 10B, Table 5). Within this manuscript, the author meticulously examines the nexus between the ethical paradigm of collaborative decision-making and AI mechanisms proffering therapeutic suggestions.

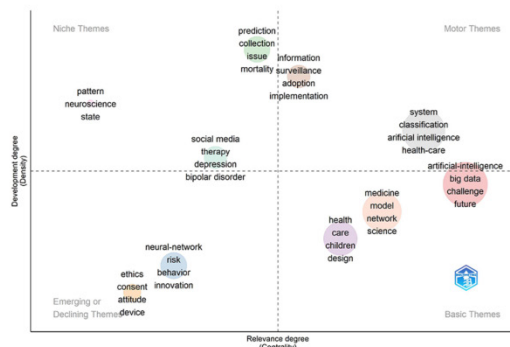
Figure 10C is a Pennant diagram for MCDougall RJ's document published in 2019. This document is positioned at the far right as one of the vertices of the triangle (given seed). In

the cognitive effect coordinate, the closer the literature is to this point, the greater the association with it; in the ease of processing coordinate, the smaller the coordinate value, the greater the association with this document or given seed. This forms a two-dimensional diagram of the collaboration relationship, indicating the strength of associations: strongly associated main bodies (green), moderately associated main bodies (orange) and weakly associated main bodies (yellow). This indicates that this

## A Top 20 Keywords with the Strongest Citation Bursts

| Keywords                  | Year | Strength | Begin | End  | 2013 - 2023 |
|---------------------------|------|----------|-------|------|-------------|
| artificial neural network | 2017 | 1.64     | 2017  | 2019 |             |
| ethical issue             | 2017 | 1.6      | 2017  | 2020 |             |
| image                     | 2018 | 1.71     | 2018  | 2019 |             |
| dementia                  | 2019 | 1.97     | 2019  | 2020 |             |
| intelligence              | 2019 | 1.85     | 2019  | 2019 |             |
| information               | 2019 | 1.7      | 2019  | 2019 |             |
| autonomous vehicle        | 2019 | 1.47     | 2019  | 2020 |             |
| burden                    | 2019 | 1.23     | 2019  | 2019 |             |
| health policy             | 2020 | 1.71     | 2020  | 2020 |             |
| deep                      | 2020 | 1.71     | 2020  | 2020 |             |
| robot                     | 2013 | 1.59     | 2020  | 2021 |             |
| clinical ethics           | 2021 | 2.04     | 2021  | 2021 |             |
| mental health             | 2014 | 1.36     | 2021  | 2021 |             |
| mobile phone              | 2021 | 1.23     | 2021  | 2021 |             |
| information technology    | 2021 | 1.23     | 2021  | 2021 |             |
| social media              | 2021 | 2.22     | 2022  | 2023 |             |
| decision making           | 2022 | 1.8      | 2022  | 2023 |             |
| network                   | 2022 | 1.8      | 2022  | 2023 |             |
| design                    | 2022 | 1.44     | 2022  | 2023 |             |
| segmentation              | 2022 | 1.44     | 2022  | 2023 |             |

## B



**Figure 12:** (A) Graphical depiction highlighting keywords with the most pronounced citation surges. (B) The thematic map of AI ethics researches.

paper has a generally strong correlation with other papers in the field of AI medical ethics research, collectively forming a co-cited references network group and taking on the role of a central node.

### Analysis of keywords

The study encompassed a cumulative of 294 author-specific keywords. Based on the chronological order of keyword occurrence, different colors are used to mark all the keywords and their collaborative network connections. From the Network map of keywords related to AI ethics researches (Figure 11A), it can be observed that "artificial intelligence" is positioned at the center of the visualization map. Additionally, keywords such as "machine learning," "big data," and "artificial neural network" have centralities greater than 0.1, indicating their relevance and significance as hot topics in the research area. It is worth noting that many keywords, such as "ethics issue," "privacy," and "medical ethics," appear more frequently after 2018, suggesting the growing research interest in the field in recent years. This observation is further supported by the thematic evolution map (Figure 11F).

The clustering network map consists of 10 distinct clusters and the Q value (0.854) and S value (0.921) confirm the rationality of the network (Figure 11B). The largest cluster is #0 managerial accounting, followed by #1 information technology and #2 interpretability. The four keyword nodes with centralities greater than 0.1, mentioned earlier, are found in these three major clusters as well. To delve deeper into the keywords associated with AI in medical ethics, a timeline view analysis was undertaken (Figure 11C). The research focus in this field has shifted from initial aspects such as "diagnosis" and "health system" to topics like "ethical issue," "performance," and "clinical ethics" after the research boom in 2015.

Burst detection algorithm is a potent analytical instrument adept at identifying pivotal junctures in the popularity of keywords or citations during specific time periods. The algorithm results in Figure 12, where the blue line symbolizes the temporal segments and the red line denotes the duration of surge. The earliest keyword to appear is "robot," which emerged in 2013 but only experienced a burst in 2020, lasting for only one year. The longest burst duration is observed for "ethical issue," which began in 2017 and lasted until 2020. The keyword with the highest burst intensity is "social media" (2.22), which started bursting in 2021 and continues to the present. From the keyword burst map, it can be seen that the current research focus and hot topics in the field of AI in medical ethics will continue to center around the trends of "social media," "decision making," "network," "design," and "segmentation." This indicates that the field of AI in medical ethics still holds value for sustained and in-depth research. Looking at the thematic map (Figure 12B), it is evident that research on AI has relatively concentrated and stable thematic directions (motor themes) with a balanced distribution of themes in the niche themes and emerging and declining themes quadrants. It is worth noting that research on medical ethics is mainly distributed in the emerging and declining themes quadrant, indicating that although it is a burst research hotspot, the research on this topic is not stable, which raises uncertainties regarding its sustainability. In summary, the field of AI in medical ethics requires more attention and efforts from researchers to delve deeper into its various aspects.

### DISCUSSION

In the current era characterized by exponential information growth, the accumulation of research outputs in a specific domain presents a significant challenge to researchers who strive to sustain a keen perceptiveness towards focal areas of research

and stay abreast of the latest advancements. Diverging from systematic reviews or meta-analyses, bibliometric analysis brings forth the advantages of encapsulating the dynamic developments within a specific research domain and offering insights into research hotspots.<sup>28</sup> Employing bibliometric methodologies, this study systematically examines the evolution and trends within the field of AI in medical ethics from 2000 to 2023, while also forecasting future research hotspots.

This study involved a total of 142 authors, 456 cited authors, 60 countries/regions, 200 research institutions, 219 academic journals, 327 papers (241 original articles and 86 reviews), 396 references and 294 author keywords. Thanks to the advancements in deep learning and the advent of the "big data" era, the field of study has gradually gained attention. The related literature has experienced a period of stagnation before entering a phase of rapid development since 2015. In terms of country/region distribution, the United States (89, 27%) is the top contributor of published papers, followed by the United Kingdom (57, 17%) and Germany (38, 12%). These three countries hold significant and leading positions in research within this field. China ranks fourth (32, 10%), with a comparable number of articles to Germany, but lags behind in total citation, average citation per article and h-index, even ranking lower in the top 10. This indicates that China's research in this field currently emphasizes quantity over quality and lacks high-quality research and output. Regarding institutional distribution, the University of Oxford has the highest publication volume (15), but its centrality is relatively low. However, it possesses a high h-index, suggesting that the institution has limited collaboration with other institutions but maintains high research quality. Most other institutions have centrality values below 0.1, indicating a decentralized research landscape within this field. In terms of authorship, Floridi Luciano ranks first in both publication volume and co-citation. Floridi Luciano is an important author in this field. It is worth mentioning that authors from the United States, the United Kingdom and Germany are frequently found among the top ranks in both publication volume and co-citation. When examining journals and co-cited journals, the utilization of dual-map overlap reveals a growing trend towards interdisciplinary attributes and diversification within this field as it progresses. These research findings suggest that to address the current development bottleneck and further explore the value of this field, it is crucial to break down barriers between countries and regions. Accelerate international collaboration and promoting direct exchanges among institutions and authors from different countries will enrich the network cooperation framework. This approach will help centralize research centers while preserving their individual characteristics and ultimately drive the overall development of the field.

Manuscripts with extensive citations often exemplify superior research caliber, showcasing pronounced innovation and marked influence within a distinct domain. The article with the

highest local citation is "Computer knows best? The need for value-flexibility in medical AI" by MCDOUGALL RJ, published in the *Journal of Medical Ethics* in 2019. Another highly cited article is "Principles alone cannot guarantee ethical AI" by MITTELSTADT B, published in *Nature Machine Intelligence* in the same year. These papers have attracted significant attention and citations, indicating their importance and influence in the field of AI ethics in medicine. MCDOUGALL RJ's research specifically scrutinized the nexus between the ethical tenet of collaborative decision-making and AI mechanisms proffering therapeutic advisories, with a specific focus on IBM's Watson for Oncology.<sup>20</sup> On the other hand, MITTELSTADT B argues that there are four crucial differences in AI development compared to medicine, which indicate that consensus around high-level principles should not be prematurely celebrated. These differences include the absence of (1) Shared objectives and trust-based responsibilities, (2) Historical context and professional standards, (3) Proven methodologies for transposing principles into actionable measures and (4) Sturdy legal and professional oversight structures.<sup>21</sup> These distinctions highlight the existence of deep political and normative disagreements in the field, which should not be overlooked when discussing the ethical implications of AI in medicine. The most co-cited article is "The global landscape of AI ethics guidelines"<sup>22</sup> published in *Nature Machine Intelligence* in 2019 by Jobin A and colleagues, with a local citation count of 38. The results elucidated in this manuscript signal the advent of a worldwide consensus encompassing five ethical tenets: lucidity, equity, harm avoidance, duty and confidentiality. However, there is significant variation in the interpretation of these principles, the reasons for their importance, the specific issues, domains, or actors to which they apply and the recommended approaches for their implementation. In other words, while there is a broad agreement on the core ethical principles, there are notable differences in their understanding and application across different contexts and perspectives. This indicates that current research on ethical AI revolves around and develops based on the ethicality, principles, standardization and characteristics of AI applications. These aspects serve as the central focus for research and provide the foundation for further exploration and development in the field.

Co-occurrence analysis of keywords is a commonly used method in bibliometrics to identify popular research topics. It reflects the changes in research themes within a field and helps to grasp research hotspots.<sup>29</sup> In the context of this study, "artificial intelligence," "machine learning," and "big data" are the most frequently occurring keywords, aligning with the main focus of the research. Burst detection, on the other hand, is an effective method to identify sudden increases in citations or keyword popularity within a specific time frame, which can help identify hot topics or themes. The earliest keyword to appear was "robot" in 2013. As the field of robotics evolved over the years, people began to recognize both its benefits and the ethical concerns

it raised. This led to the emergence of more specific research topics and discussions, such as "ethical issue," "decision making," "privacy," and "clinical ethics." Although the field is currently in a phase of rapid development, it is still some distance away from a state of vibrant discourse and diverse perspectives. The depth of research and the prospects for sustainable development in this area remain somewhat unclear.

During the period from 2000 to 2018, the keywords "diagnosis," "management," and "health system" emerged. This indicates that this phase was characterized by a research focus on the functional aspects of "robot" in various professional domains. Reggia JA offers three conclusions regarding the current state of the field. First of all, computational modeling has emerged as a valuable and widely accepted methodology for studying consciousness scientifically. Secondly, prevailing computational frameworks have adeptly encapsulated an array of neurobiological, cognitive and behavioral facets tied to conscious data processing via machine emulations. In the final analysis, no current methodology in the realm of artificial consciousness has compellingly exhibited phenomenal machine awareness or proffered irrefutable proof that the realization of artificial phenomenal consciousness is within reach.<sup>30</sup> To analyze new cases, it is beneficial to utilize extensionally defined principles and refer to relevant past cases. In order to explore this phenomenon computationally, McLaren BM conducted an analysis of an expansive dataset encompassing professional ethics instances was collated, leading to the formulation of a computational blueprint termed SIROCCO. This model serves as a mechanism devised for the extraction of guiding principles and historical instances, aiding in the investigation of ethical principles and cases.<sup>31,32</sup>

Since 2018, research on ethical AI has become more aligned with the current era's background and demands. Keywords such as "deep learning," "AI ethics," "medical ethics," "technology," "model," and "design" have emerged. This shift indicates that the field has transitioned from basic functional research to studying the algorithms and ethical implications of AI. Machine learning, as a subfield of artificial intelligence, systematically utilizes algorithms to uncover potential relationships among data and information.<sup>32</sup> Currently, Convolutional Neural Networks (CNNs) and deep learning techniques are widely employed in various fields.<sup>33</sup> The popularity of these methods can be attributed, in part, to the seminal work by Breiman L in 2001 on Random Forests. Breiman's research introduced a machine learning algorithm that demonstrated enhanced resilience to noise, thereby providing a robust foundation for subsequent studies in the field. Indeed, these algorithms are being applied in various types of research.<sup>34-36</sup> Following the development of mature AI technologies, the discussion on their ethical implications has emerged. Morley J *et al.* discovered that ethical concerns in the context of AI can manifest in three main categories. Firstly, there are epistemic issues, which pertain to problems arising from

misguided, inconclusive, or inscrutable evidence. Secondly, normative issues arise in relation to unfair outcomes and the transformative effects of AI. Lastly, traceability issues refer to the challenges associated with tracking and attributing responsibility in AI systems. Furthermore, these ethical issues are found to operate at multiple levels of abstraction, including the individual, interpersonal, group, institutional and societal or sectoral levels.<sup>15</sup> These discussions have become more in-depth and specific, encompassing a range of controversies. In the realm of medical ethics, Keskinbora KH highlights the inherent possibility of errors and unintended consequences within AI algorithms, which may lead to inequitable outcomes based on economic and racial distinctions. Therefore, it is of utmost importance to establish robust mechanisms for monitoring technological advancements and implementing preventive and precautionary measures. These measures are essential to protect the rights of individuals involved in AI applications, safeguarding them against any form of direct or indirect coercion.<sup>37</sup> In addition to that, there are also controversies in similar aspects.<sup>38,39</sup> In line with current needs, there is a growing demand for the application of AI in various technological domains. Ambient Intelligence (AmI) refers to the integration and utilization of artificial intelligence within everyday environments providing seamless and intuitive support through an imperceptible user interface. This technology is applied across various contexts such as autonomous vehicles, smart homes, industrial domains and healthcare settings, collectively known as Ambient Assistive Living.<sup>40</sup> Furthermore, in the clinical realm, the application of AmI has primarily focused on diagnosis and predictive analytics for outcomes assessment.<sup>41</sup> Chan *et al.*<sup>42</sup> conducted a comprehensive content analysis, shedding light on the utilization of Artificial Intelligence (AI) in the education sector, aiming to identify prevailing research trends and challenges within this domain. In the context of governmental employment of AI and its implications for citizens' privacy, Reddy *et al.*<sup>43</sup> delineated eight key areas encompassing human behavior prediction, intelligent decision-making processes, decision automation, digital surveillance, data privacy legislation and regulation, as well as the associated risks of behavioral modification.

The results of bibliometric research are objective and accurate, providing comprehensive assistance to scholars and researchers engaged in the field. Given the increasingly important role and position of AI technology in today's society, as well as its significant advantages in various industries, there is no doubt that research in the field of ethical AI will become a current research hotspot and direction, in parallel with the widespread use of AI.

## LIMITATIONS

This study has several limitations that warrant discussion. First, this research exclusively selected the Web of Science Core Collection (WoSCC) as the database. This choice may



have led to the omission of relevant papers indexed in other databases, potentially affecting the comprehensiveness of our analysis.<sup>44</sup> Secondly, the strict search strategy and restriction to English-language publications may result in some data loss, as non-English studies were excluded. Lastly, the dynamic nature of database updates means that recent high-quality articles might not have been captured within the timeframe of this study, possibly impacting the currency of our findings.<sup>45</sup>

## CONCLUSION

This study conducted an in-depth analysis of 327 published papers on ethical AI using bibliometric analysis. The results indicate that the research center in the field of ethical AI is currently experiencing stable and rapid development. The research center is mainly concentrated in North America, with a focus on the United States, United Kingdom and Germany. Furthermore, the prevailing focal areas of research and forthcoming trajectories of inquiry will continue to focus on technological applications, networks and decision-making. By strengthening international cooperation, the field can achieve comprehensive and systematic development, thereby opening up new frontiers and research hotspots.

## FUNDING

This work was financially supported by 2018 National Major Social Science Foundation of China (18ZDA322) and 2023 Postgraduate Research and Practice Innovation Program of Jiangsu Province (KYCX23\_2007).

## CONFLICT OF INTEREST

The sponsors had no role in the design, execution, interpretation, or writing of the study.

## ABBREVIATIONS

**AI:** Artificial Intelligence; **WoS:** The Web of Science; **TS:** Term Search; **CNNs:** Convolutional Neural Networks; **AmI:** Ambient Intelligence; **WoSCC:** The Web of Science Core Collection; **LCS:** Local cCitation Score; **GCS:** Global Citation Score; **SCP:** Single Country Publications; **MCP:** Multiple Country Publications; **TC:** Total Citations.

## AUTHOR CONTRIBUTIONS

Yuan Qihui: Conceptualization, Writing - Original Draft. Chen Tingxuan: Resources, Data Curation, Software. Yang Qing: Writing - Original Draft, Formal analysis. Zhang Zongming: Writing - Review and Editing, Funding acquisition.

## SUMMARY

This study meticulously investigated the evolution and main focus areas of ethical AI research. Using the Web of Science Core Collection, it analyzed 327 articles and applied tools like CiteSpace, VOSviewer and the R package to create bibliometric maps and networks. Key findings included the rapid increase in publications from 2015, significant contributions from the USA and the University of Oxford and prevalent topics like privacy and medical ethics. The analysis highlighted the importance of ethical considerations in AI applications and suggested a need for enhanced international and interdisciplinary collaboration.

## REFERENCES

- Balyen L, Peto T. Promising artificial intelligence-machine learning-deep learning algorithms in ophthalmology. *Asia Pac J Ophthalmol.* 2019;8(3):264-72.
- Buruk B, Ekmekci PE, Arda B. A critical perspective on guidelines for responsible and trustworthy artificial intelligence. *Med Health Care Philos.* 2020;23(3):387-99. doi: 10.1007/s11019-020-09948-1.
- Prakash S, Balaji JN, Joshi A, Surapaneni KM. Ethical conundrums in the application of artificial intelligence (AI) in healthcare-A scoping review of reviews. *J Pers Med.* 2022;12(11):1914. doi: 10.3390/jpm12111914.
- Lee D, Yoon SN. Application of artificial intelligence-based technologies in the healthcare industry: opportunities and challenges. *IJERPH.* 2021;18(1):271. doi: 10.3390/ijerph18010271.
- Köbis L, Mehner C. Ethical questions raised by AI-supported mentoring in higher education. *Front Artif Intell.* 2021;4:624050. doi: 10.3389/frai.2021.624050.
- Walker R, Dillard-Wright J, Iradukunda F. Algorithmic bias in artificial intelligence is a problem-And the root issue is power. *Nurs Outlook.* 2023;71(5):102023. doi: 10.1016/j.outlook.2023.102023.
- Adlassnig KP. Artificial-intelligence-augmented systems. *Artif Intell Med.* 2002;24(1):1-4. doi: 10.1016/S0933-3657(01)00102-6.
- Pashkov VM, Harkusha AO, Harkusha YO. Artificial intelligence in medical practice: regulative issues and perspectives. *Wiad Lek.* 2020; 73(12)(12 cz 2):2722-7. doi: 10.36740/WLek202012204.
- Yeung AW, Tzvetkov NT, Balacheva AA, Georgieva MG, Gan RY, Jozwik A, *et al.* Lignans: quantitative analysis of the research literature. *Front Pharmacol.* 2020;11:37. doi: 10.3389/fphar.2020.00037.
- Abdelwahab SI, Taha MM, Aldhahi MI. Comprehensive analysis of research related to rehabilitation and COVID-19, hotspots, mapping, thematic evolution, trending topics and future directions. *Eur J Med Res.* 2023;28(1):434. doi: 10.1186/s40001-023-01402-1.
- Yang K, Hu Y, Qi H. Digital health literacy: bibliometric analysis. *J Med Internet Res.* 2022;24(7):e35816. doi: 10.2196/35816.
- Chen C, Song M. Visualizing a field of research: A methodology of systematic scientometric reviews. *PLOS ONE.* 2019;14(10):e0223994. doi: 10.1371/journal.pone.0223994.
- Wu H, Cheng K, Tong L, Wang Y, Yang W, Sun Z. Knowledge structure and emerging trends on osteonecrosis of the femoral head: a bibliometric and visualized study. *J Orthop Surg Res.* 2022;17(1):194. doi: 10.1186/s13018-022-03068-7.
- Chen C. An information-theoretic view of visual analytics. *Comput Graph.* 2008;28(1):18-23. Srivastava R, Srivastava S. Bibliometric analysis of Indian journal of palliative care from 1995 to 2022 using the VOSviewer and Bibliometrix software. *Indian J Palliat Car.* 2022;28(4):338-53. doi: 10.25259/IJPC\_30\_2022.
- Morley J, Machado CC, Burr C, Cows J, Joshi I, Taddeo M, *et al.* The ethics of AI in health care: A mapping review. *Soc Sci Med.* 2020;260:113172. doi: 10.1016/j.socscimed.2020.113172.
- Ioannidis JP, Baas J, Klavans R, Boyack KW. A standardized citation metrics author database annotated for scientific field. *PLOS Biol.* 2019;17(8):e3000384. doi: 10.1371/journal.pbio.3000384.
- Xiong M, Xu Y, Zhao Y, He S, Zhu Q, Wu Y, *et al.* Quantitative analysis of artificial intelligence on liver cancer: A bibliometric analysis. *Front Oncol.* 2023;13:990306. doi: 10.3389/fonc.2023.990306.
- Zhang G, Song J, Feng Z, Zhao W, Huang P, Liu L, *et al.* Artificial intelligence applicated in gastric cancer: A bibliometric and visual analysis via CiteSpace. *Front Oncol.* 2022;12:1075974. doi: 10.3389/fonc.2022.1075974.
- McDougall RJ. Computer knows best? The need for value-flexibility in medical AI. *J Med Ethics.* 2019;45(3):156-60. doi: 10.1136/medethics-2018-105118.
- Robinson DB, Hopkins L, Brown C, Abdelrahman T, Powell AG, Egan RJ, *et al.* Relative value of adapted novel bibliometrics in evaluating surgical

- academic impact and reach. *World J Surg.* 2019;43(4):967-72. doi: 10.1007/s00268-018-04893-w.
21. Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, *et al.* AI4People-An ethical framework for a good AI society: opportunities, risks, principles and recommendations. *Mind Mach.* 2018;28(4):689-707. doi: 10.1007/s11023-018-9482-5.
  22. Wu J, Ji S, Niu P, Zhang B, Shao D, Li Y, *et al.* Knowledge mapping of syringomyelia from 2003 to 2022: A bibliometric analysis. *J Clin Neurosci.* 2023;110:63-70. doi: 10.1016/j.jocn.2023.01.004.
  23. Xie X, Zhang P, Ran C, Liu L, Hu J, Lei P, *et al.* Global research status and hotspots of radiotherapy for prostate cancer: a bibliometric analysis based on Web of Science from 2010-2022. *Front Oncol.* 2023;13:1135052. doi: 10.3389/fonc.2023.1135052.
  24. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* 2019;25(1):44-56. doi: 10.1038/s41591-018-0300-7.
  25. Luxton DD. Recommendations for the ethical use and design of artificial intelligent care providers. *Artif Intell Med.* 2014;62(1):1-10. doi: 10.1016/j.artmed.2014.06.004.
  26. Wu H, Tong L, Wang Y, Yan H, Sun Z. Bibliometric analysis of global research trends on ultrasound microbubble: A quickly developing field. *Front Pharmacol.* 2021;12:646626. doi: 10.3389/fphar.2021.646626.
  27. Zeng J, Chu H, Lu Y, Xiao X, Lu L, Li J, *et al.* Research status and hotspots in the surgical treatment of tremor in Parkinson's disease from 2002 to 2022: a bibliometric and visualization analysis. *Front Aging Neurosci.* 2023;15:1157443. doi: 10.3389/fnagi.2023.1157443.
  28. Reggia JA. The rise of machine consciousness: studying consciousness with computational models. *Neural Netw.* 2013;44:112-31. doi: 10.1016/j.neunet.2013.03.011.
  29. Rodriguez-Ruiz A, Lâng K, Gubern-Merida A, Broeders M, Gennaro G, Clauser P, *et al.* Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists. *Jnci-J N Atl Cancer.* 2019;111(9):916-22. doi: 10.1093/jnci/djy222.
  30. Weber S, Wyszynski M, Godefroid M, Plattfaut R, Niehaves B. How do medical professionals make sense (or not) of AI? A social-media-based computational grounded theory study and an online survey. *Comput Struct Biotechnol.* 2024;24:146-59. doi: 10.1016/j.csbj.2024.02.009.
  31. Badillo S, Banfai B, Birzele F, Davydov II, Hutchinson L, Kam-Thong T, *et al.* An introduction to machine learning. *Clin Pharmacol Ther.* 2020;107(4):871-85. doi: 10.1002/cpt.1796.
  32. Hwang DK, Hsu CC, Chang KJ, Chao D, Sun CH, Jheng YC, *et al.* Artificial intelligence-based decision-making for age-related macular degeneration. *Theranostics.* 2019;9(1):232-45. doi: 10.7150/thno.28447.
  33. Lawrence S, Giles CL, Tsoi AC, Back AD. Face recognition: a convolutional neural-network approach. *IEEE Trans Neural Netw.* 1997;8(1):98-113. doi: 10.1109/72.554195.
  34. Keskinbora KH. Medical ethics considerations on artificial intelligence. *J Clin Neurosci.* 2019;64:277-82. doi: 10.1016/j.jocn.2019.03.001.
  35. Peiffer-Smadja N, Rawson TM, Ahmad R, Buchard A, Georgiou P, Lescure FX, *et al.* Machine learning for clinical decision support in infectious diseases: a narrative review of current applications. *Clin Microbiol Infect.* 2020;26(5):584-95. doi: 10.1016/j.cmi.2019.09.009.
  36. Fosso WS, Queiroz MM. Responsible artificial intelligence as a secret ingredient for digital health: bibliometric analysis, insights and research directions; 2021. *Inform Syst Front.* p. 1-16. Available from: <https://pubmed.ncbi.nlm.nih.gov/34025210/>.
  37. Martinez-Martin N, Kreitmair K. Ethical issues for direct-to-consumer digital psychotherapy apps: addressing accountability, data protection and consent. *JMIR Ment Health.* 2018;5(2):e32. doi: 10.2196/mental.9423.
  38. Biller-Andorno N, Ferrario A, Joebges S, Krones T, Massini F, Barth P *et al.* AI support for ethical decision-making around resuscitation: proceed with care. *J Med Ethics.* 2022;48(3):175-83. doi: 10.1136/medethics-2020-106786.
  39. Martinez-Martin N, Luo Z, Kaushal A, Adeli E, Haque A, Kelly SS, *et al.* Ethical issues in using ambient intelligence in health-care settings. *Lancet Digit Health.* 2021;3(2):e115-23. doi: 10.1016/S2589-7500(20)30275-2.
  40. Ma D, Yang B, Guan B, Song L, Liu Q, Fan Y, *et al.* A bibliometric analysis of pyroptosis from 2001 to 2021. *Front Immunol.* 2021;12:731933. doi: 10.3389/fimmu.2021.731933.
  41. Chan PY, Tay A, Chen D, De Freitas M, Millet C, Nguyen-Duc T, *et al.* Ambient intelligence-based monitoring of staff and patient activity in the intensive care unit. *Aust Crit Care.* 2023;36(1):92-8. doi: 10.1016/j.aucc.2022.08.011.
  42. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. *J Am Med Inform Assoc.* 2020;27(3):491-7. doi: 10.1093/jamia/ocz192.
  43. Ma D, Yang B, Guan B, Song L, Liu Q, Fan Y, *et al.* A bibliometric analysis of pyroptosis from 2001 to 2021. *Front Immunol.* 2021;12:731933. doi: 10.3389/fimmu.2021.731933.
  44. Cheek J, Garnham B, Quan J. What's in a number? Issues in providing evidence of impact and quality of research(ers). *Qual Health Res.* 2006;16(3):423-35. doi: 10.1177/1049732305285701.

**Cite this article:** Yuan Q, Chen T, Yang Q, Zhang Z. Unveiling Ethical AI: An In-Depth Bibliometric and Visual Exploration. *Indian J of Pharmaceutical Education and Research.* 2024;58(4):1084-101.