

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/350582963>

Ethics and privacy of artificial intelligence: Understandings from bibliometrics

Article in Knowledge-Based Systems · April 2021

DOI: 10.1016/j.knosys.2021.106994

CITATIONS

41

READS

3,108

5 authors, including:



Yi Zhang

University of Technology Sydney

126 PUBLICATIONS 2,529 CITATIONS

SEE PROFILE



Mengjia Wu

University of Technology Sydney

36 PUBLICATIONS 373 CITATIONS

SEE PROFILE



George Yijun Tian

University of Technology Sydney

11 PUBLICATIONS 81 CITATIONS

SEE PROFILE



Jie Lu

University of Technology Sydney

786 PUBLICATIONS 21,521 CITATIONS

SEE PROFILE

Please cite as: Zhang, Y., Wu, M., Tian, G. Y., Zhang, G. & Lu, J. 2021, Ethics and privacy of artificial intelligence: Understandings from bibliometrics, *Knowledge-based Systems*, DOI: 10.1016/j.knosys.2021.106994

Ethics and privacy of artificial intelligence: Understandings from bibliometrics

Yi Zhang¹, Mengjia Wu¹, George Yijun Tian², Guangquan Zhang¹, Jie Lu¹

¹Australian Artificial Intelligence Institute, Faculty of Engineering and Information Technology,
University of Technology Sydney, Australia

²Faculty of Law, University of Technology Sydney, Australia

Email: yi.zhang@uts.edu.au; mengjia.wu@student.uts.edu.au; yijun.tian@uts.edu.au;
guangquan.zhang@uts.edu.au; jie.lu@uts.edu.au

ORCID: 0000-0002-7731-0301 (Yi Zhang); 0000-0003-3956-7808 (Mengjia Wu); 0000-0003-4472-5428 (George Yijun Tian); 0000-0003-3960-0583 (Guangquan Zhang); 0000-0003-0690-4732 (Jie Lu).

Abstract

Artificial intelligence (AI) and its broad applications are disruptively transforming the daily lives of human beings and a discussion of the ethical and privacy issues surrounding AI is a topic of growing interest, not only among academics but also the general public. This review identifies the key entities (i.e., leading research institutions and their affiliated countries/regions, core research journals, and communities) that contribute to the research on the ethical and privacy issues in relation to AI and their intersections using co-occurrence analysis. Topic analyses profile the topical landscape of AI ethics using a topical hierarchical tree and the changing interest of society in AI ethics over time through scientific evolutionary pathways. We also paired 15 selected AI techniques with 17 major ethical issues and identify emerging ethical issues from a core set of the most recent articles published in *Nature*, *Science*, and *Proceedings of the National Science Academy of the United States*. These insights, bridging the knowledge base of AI techniques and ethical issues in the literature, are of interest to the AI community and audiences in science policy, technology management, and public administration.

Keywords

Artificial intelligence; Ethics; Privacy; Bibliometrics; Topic analysis.

Highlights

- Articles on AI ethics cover 199 of the 254 Web of Science Categories, indicating a broad interest from the academia.
- Research communities of computer science, business and management, medical science and law are playing a leading role on studies of AI ethics.
- USA, UK, and China make the major contribution to AI ethics, with a relatively high level of domestic collaborations.
- Key AI techniques raise ethical concerns, such as fairness, accountability, data privacy, responsibility, liability, and crimes.

1. Introduction

A Pandora's box of artificial intelligence (AI) has been opened and these disruptive technologies are transforming the daily lives of human beings in relation to new ways of thinking and behavioral patterns, with enhanced capabilities and efficiency. There are many examples of AI applications in use today, such as smart homes [1] smart farming [2], precision medicine [3] and healthcare surveillance systems [4]. The ethical and privacy issues surrounding the use of AI have been a topic of growing interest among diverse communities. For example, the general public has expressed concern about the impact of the increased use of robots on unemployment and inequality [5], social scientists have raised deep privacy concerns related to surveillance systems [6], and limited regulation of social media has raised debate with technical giants on the abuse of private data¹. Despite these concerns, the AI community stands behind the efficiency and robustness of their AI models and there is an urgent need to guide the research community to understand these ethical and privacy challenges.

Bibliometrics, which is a set of approaches for analyzing scientific documents (e.g., research articles, patents, and academic proposals), has been widely used as a tool for science, technology and innovation studies [7], such as identifying technological topics [8], discovering latent relationships [9], and predicting potential future trends [10]. Recently, AI has received recognition in bibliometrics as an emerging topic for empirical investigation [11, 12]. These investigations either align with the interest in technology management (e.g., using AI as a representative case in digital transformation) or emphasize its role in examining the reliability of the proposed methods. However, from a practical perspective, a bibliometric guide which summarizes ideas, assumptions, and debate in the literature would bring significant benefits to the AI community, not only by highlighting the ethical and privacy concerns raised by the public but also by identifying the potential conflicts between AI techniques and these issues of concern.

To address these concerns, this paper reports on a bibliometric study to comprehensively profile the key ethical and privacy issues discussed in the research articles and to trace how such issues have changed over the past few decades. We integrated a set of intelligent bibliometric approaches within a framework for diverse analyses. To identify the key entities, i.e., the leading research institutions and their affiliated countries and regions, and the core research journals and their behind research communities, which report the ethical and privacy issues surrounding AI, we used co-occurrence statistics with diverse bibliographical indicators (e.g., authors, affiliations, and sources). With specific foci in topic analysis, we initially retrieved terms from the combined titles and abstracts of collected articles and used a term clumping process [13] to remove noisy terms and consolidate technical synonyms. In parallel, we represented each word in the combined field with titles and abstracts as a vector using the Word2Vec model [14] and combined the word vectors into term vectors by matching the core terms refined in the term clumping process. We answered the

¹ More information can be found on the website: <https://www.bbc.com/news/business-49099364>

questions as to what is the topical landscape and how have these topics evolved over time, using an approach of scientific evolutionary pathways [15]. We also targeted a core set of articles published in three world-leading multi-disciplinary journals, namely *Science*, *Nature*, and *Proceedings of the National Academy of Sciences* (PNAS) of the United States of America, and identified cutting-edge issues that might either focus attention on emergent ethical and privacy issues in the current AI age or lead to novel developments in AI models to address any potential negative impacts. We anticipate that the empirical insights identified in this study will motivate the AI community to extensively and comprehensively discuss the ethical and privacy issues surrounding AI and will guide the implementation of AI in line with an ethical framework.

The rest of this paper is organized as follows: Section 2 presents a review of the related work on AI ethics, privacy, and bibliometrics; Section 3 introduces the data and methodologies used in this study; Section 4 presents the results, and our key findings and Section 5 concludes the study and suggests future research directions.

2. Related work

In this section, we review the current debate on the ethical and privacy issues surrounding AI and then briefly introduce the bibliometrics and topic analysis used in this study.

2.1. Ethics, ethical dilemma, and AI ethics

In philosophy, ethics describes “what is good for the individual and for society”, as well as the essence of “duties that people owe themselves and one another” [16], while ethical dilemma refers to certain ethical problems can be extremely complicated and the challenges they bring cannot be easily solved. Ever-improving technologies bring along with multiple advantages to human society, but they may also “generate downside risks and challenges, including more complicated ethical dilemma². This is true with AI technologies.

With the rapid growth in AI techniques in recent decades, there has been increasing controversy over the impact of AI on the daily lives of human beings, for example, the potential for robots to replace human labor [17], the accountability and accident risk of driverless vehicles [18], the self-awareness and behavior autonomy of robotics [19], and possible fraud caused by deep-fake videos and photos [20]. Such concerns in relation to the ethics around AI has attracted attention from global federal governments and corporations, in particular, tech giants such as Google and SAP, when those corporations are willing to form national and industrial committee to formulate AI ethics guidelines [21]. An increasing number of international organizations have also started to take actions to address the ethical challenges brought by AI technology. As one of the most recent developments, the United Nations Educational, Scientific and Cultural Organization (UNESCO) has issued its first draft of *Recommendation on the Ethics of*

² the United Nations Educational, Scientific and Cultural Organization, *Elaboration of a Recommendation on the ethics of artificial intelligence* at <https://en.unesco.org/artificial-intelligence/ethics>

*Artificial Intelligence (Recommendations)*³ in September 2020, which sets up ten important Principles of the Ethics of AI, including: proportionality and do no harm, safety and security, fairness and non-discrimination, sustainability, privacy, human oversight and determination, transparency and expandability, responsibly and accountability, awareness and literacy, and multi-stakeholder and adaptive governance and collaboration.

2.2. *Privacy, data privacy, and AI privacy*

Privacy, as one of ten important Principles of the Ethics of AI developed by the UNESCO, may deserve a particular attention. In legal and philosophical literature, privacy has been defined in a variety of ways, for example, privacy is “the right to be let alone”, as a component of personhood, control over personal information, and the right to secrecy [6].

Together with the big data boom and AI age, data privacy⁴ and control over personal information arguably becomes increasingly important aspects of privacy protection, and AI brings further threats to privacy protection [22]. Kerry (2020) observed that AI expands the ability to use personal information in ways that can infringe on privacy interests by bringing personal data analysis to new levels of power and speed [23].

2.3. *Ethical dilemma of AI academia*

As discussed above, ethical dilemmas may exist almost every aspect of human lives, including personal, social, and professional. Academia certainly cannot be exempted from this either. The AI community not only contributes to the development of novel AI techniques but also responds to these increasing concerns about ethical dilemmas as evidenced by the large number of review papers in academic journals and online monographies on this issue. Jobin et al. [24] conducted a meta-analysis for 84 core documents that discuss ethical principles or guidelines for AI, highlighting eleven universal principals for AI ethics: transparency, justice, fairness & equity, non-maleficence, privacy, trust, responsibility, accountability, beneficence, sustainability, and solidarity. Unfortunately, how to apply these principles and abstract guidelines to actual practice is still elusive, and the lack of reinforcement mechanisms (e.g., concrete technical methods and algorithms) to translate these principles into practice remains an unsolved challenge [25, 26].

Such endeavors motivate us to conduct a bibliometric study to systematically analyze an extensive collection of research articles on ethical dilemma of AI academia, not only to profile these principles and guidelines but also to discover their potential connections with specific AI techniques. These could provide insightful knowledge to guide the AI community in developing AI techniques and/or applying them in practice.

³ UNESCO, *Outcome document: first draft of the Recommendation on the Ethics of Artificial Intelligence* at <https://unesdoc.unesco.org/ark:/48223/pf0000373434>

⁴ Despite differences between data privacy and information privacy in the legal literature, we do not specifically distinguish the two terms in this paper.

Particularly, note that, since the above discussed principals of AI ethics, concluded by diverse parties and from diverse aspects, classified privacy as one specific element/principal of ethics, in this paper we follow this line and mainly use AI ethics as a set of these ethical and privacy issues in relation to AI.

2.4. *Bibliometrics and topic extraction*

Modern bibliometrics can be traced back to the observations of Derek Price on the patterns of scientific activities [27]. Early definitions of bibliometrics emphasize “the application of mathematics and statistical methods to books and other media of communication [28]”, involving indicators such as citation/co-citation statistics, word co-occurrence, and co-authorships [15]. The increasing diversity of practical data sources rapidly extends the scope of bibliometric data from books to a wide range of information resources in science, technology and innovation, such as research articles, patents, and academic proposals, as well as to social media data (e.g., Twitter) [29]. Information technologies, especially AI techniques, further strengthen the capabilities of bibliometrics in analyzing scalable data with enhanced efficiency, effectiveness, and robustness. In this area, some of our pilot studies are spearheading a cross-disciplinary approach that develops computational models incorporating bibliometric indicators with AI techniques, which we call *intelligent bibliometrics* [30].

Topic extraction identifies abstract topics from a collection of documents to represent the major content, using either clustering or classification algorithms [31]. Topic extraction is also of significant interest to the bibliometric community, in which citation statistics and textual elements are heavily used [8, 32]. These extracted topics represented by either a sub-collection of documents or a set of terms hold recognized capabilities in knowledge interpretation and exploration, e.g., profiling research disciplines and technological areas [7, 33], identifying latent relationships [10, 15, 34], and predicting potential future changes in either collaborative patterns or research interests [35-37]. However, regarding the characteristics of bibliometric documents and the urgent need to interpret topics in depth, we anticipate two emergent directions of topic extraction: 1) since research topics are constantly changing (e.g., cross-/inter-/multi-disciplinary interactions) rather than being stable [15], extracting topics and discovering their relationships from a dynamic perspective could be practically significant for not only the bibliometric community but also business and management studies; and 2) hierarchy is an innate structure of knowledge composition, as well as topics. Thus, profiling topics from a hierarchical dimension would provide an extensive understanding of its related knowledge base [38]. Even though it is not a new task for the computer science discipline, a balance between non-parametric solutions and explainable results is still elusive.

3. Data and methodology

3.1. Data

The Web of Science (WoS)⁵ owned by Clarivate is a well-recognized integrative platform of bibliometric data sources. Of these, the WoS All Databases covers all the WoS's subscribed resources which we used as our primary data source when considering AI ethics as an emerging topic covering both natural sciences and social sciences. Its major debates exist not only in journal articles but also in a wide range of resources (e.g., conference proceedings, and other types of research publications⁶). Our special interest is in the ethical issues surrounding AI at both the macro and micro levels. Thus, topic analyses would focus on the WoS All Databases. In addition, since the WoS Core Collection database provides a cleaned form of full bibliographical information (e.g., author affiliations, countries/regions, and forward and backward citations), we particularly focused on an analysis of the key entities that contribute to the research on AI ethics and the interactions between these entities. Comparably, the WoS All Database covers a relatively "full" collection of various types of articles in WoS, with a priority on data coverage, but the WoS Core Collection only contains journal articles collected in selective indexes (e.g., Science Citation Index), highlighting the quality of its data collection. In other words, the WoS Core Collection is a subset of the WoS All Database, with a filtered data collection.

Referring to the literature discussed in Sections 2.1-2.3 on AI ethics, together with the IEEE's *Ethically Aligned Design* [39] and the AI Ethics Principles reported by Australia's Department of Industry, Science, Energy and Resources⁷, we proposed a search string and collected data on October 14, 2020 (see Table 1)⁸. We set #1 as the full dataset for understanding the topic landscape of AI ethics, and #3 (a subset of #1) as the dataset for identifying key research entities (e.g., affiliations and communities) that contribute to the research on AI ethics. As a specific interest, we collected another subset #2 from #1, containing articles published in the three world-leading multi-disciplinary journals – i.e., *Science*, *Nature*, and *PNAS*, to discover potential emerging issues in AI ethics.

Table 1. Search strategy and data information

| Dataset | #R | Search strategy |
|---------|----|-----------------|
|---------|----|-----------------|

⁵ More information on the WoS database can be found on the website:

<https://clarivate.com/webofsciencegroup/solutions/web-of-science/>

⁶ As an example, in Nature, they have 'news & views', 'insights, reviews and perspectives', and 11 other types of contributions, but the WoS Core Collection only indexes 'research articles' while we believe AI ethics could be an appealing topic in various types of contributions.

⁷ Details of the AI Ethics Principles are listed on the website: <https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework/ai-ethics-principles>

⁸ We specifically removed the term "privacy" from the search string with the following reasons: 1) privacy is heavily related to techniques and algorithms in the research area of cybersecurity, blockchain, and internet of things, and thus a large number of technical records (but without any content on AI ethics) might be retrieved. 2) We tested the overlaps between search strings with and without the term "privacy" and noticed that those highly relevant records could be involved even in the latter search string – we considered "ethics" may mainly cover "privacy" in terms of search strategy.

| | | |
|----|------|---|
| #1 | 4375 | TS = (("artificial intelligence" OR "big data") AND ("disinform*" OR "ethic*" OR "crimin*" OR "moneti*" OR "data control*" OR "implicit trust*" OR "addiction*" OR "contestab*" OR "moral*" OR "digit* transparen*" OR "algorithm* transparen*" OR "accountabilit*" OR "liabilit*" OR "fairness*"))) Data source: WoS All Databases |
| #2 | 53 | SO = ("Science" OR "Nature" OR "Proceedings of the National Academy of Sciences") in #1 |
| #3 | 3259 | The same search string with #1 Data source: WoS Core Collection |

Note that 1) according to WoS's field tags, TS = topic, and SO = publication name; and 2) #R = the number of records.

Focusing on #1 and #2, the trends for the annual number of records in the two datasets are given in Figure 1. Before 2013, the number of records in the full dataset increased at a relatively low rate and a sudden rise after 2016 illustrates the urgent attention from the academia. Comparably, the general trend in the core dataset coincides with that of the full dataset – that is, certain isolative papers are observed before 2013 and the ‘real’ growth starts in 2014.

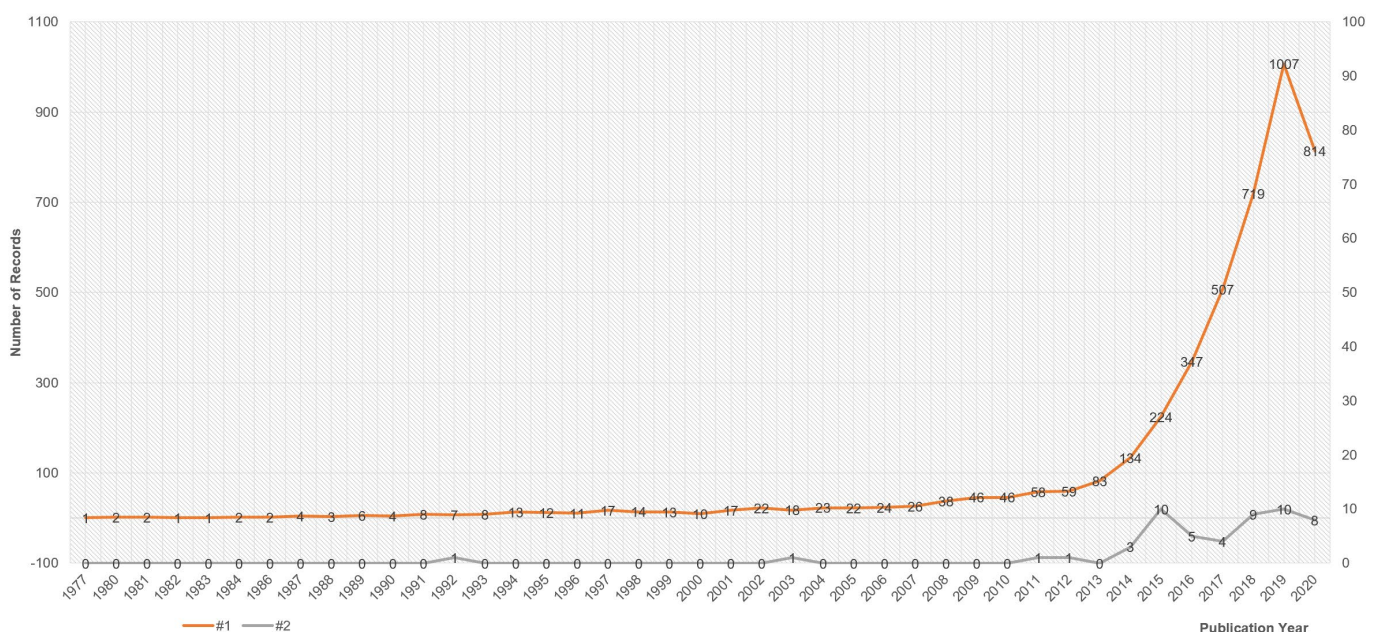


Figure 1. Trends in the annual number of records in #1 and #2 datasets

3.2. Methodology

The research framework is shown in Figure 2. We had a two-phase approach to discover insights into the ethical issues surrounding AI discussed in the research articles: phase 1 for data pre-processing and phase 2 for a systematic analysis incorporating bibliometrics with a series of analytic approaches.

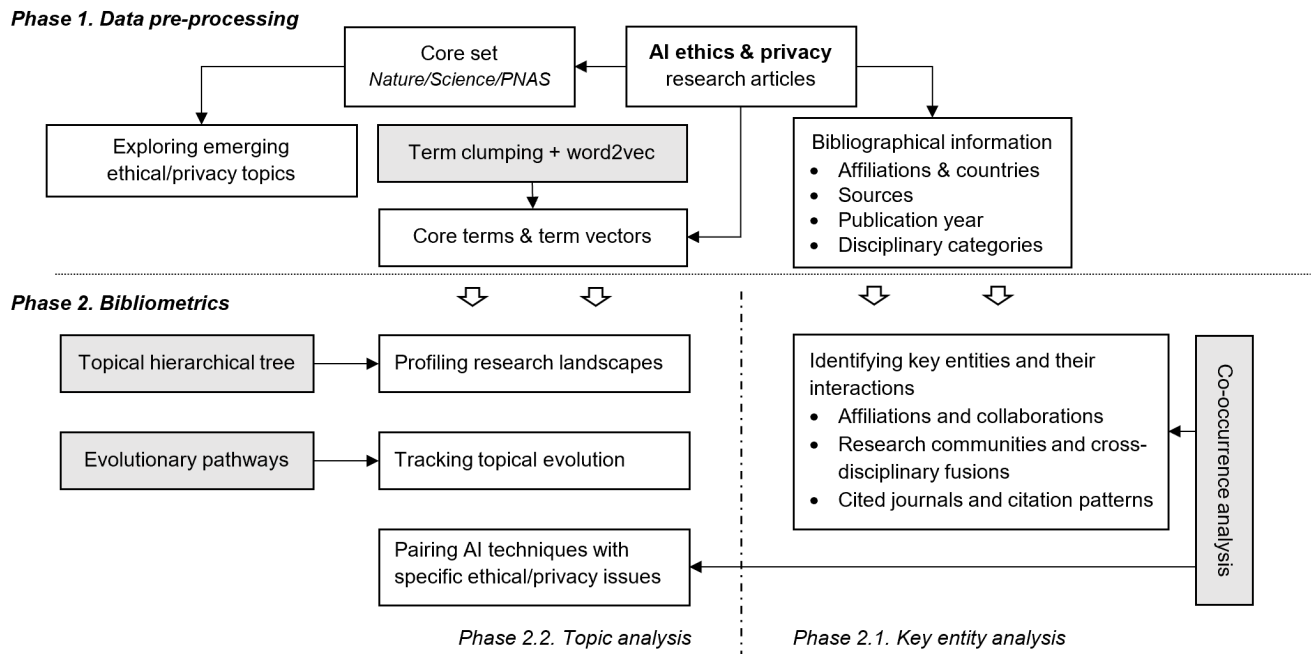


Figure 2. Research framework on understanding AI ethics and privacy

3.2.1. Phase 1 data pre-processing

In this study, we focused on two types of bibliometric indicators: 1) traditional bibliographical information including authors, affiliations, sources (e.g., journal names), and publication year; and 2) terms (e.g., words and phrases) retrieved from the titles and abstracts of research articles through natural language processing techniques. Since raw terms contain a huge number of meaningless items (e.g., pronouns, prepositions, and conjunctions) and variations (e.g., synonyms), we used a term clumping process [13] to identify a set of core terms by removing noise and consolidating variations with a set of thesauri and rules. In parallel, we applied a Word2Vec model [14] to the raw text of titles and abstracts and represent each word as a vector. We then exploited a matching function to bridge word vectors and core terms and create a vector for each core term.

Our aim was to focus on the current emerging concerns in relation to AI ethics raised by multiple research communities. We specifically collected a core set of research articles published in three top-level multidisciplinary journals, namely *Nature*, *Science*, and *PNAS*, and conducted a miniature bibliometric analysis to explore emerging issues related to AI ethics.

3.2.2. Phase 2 bibliometrics

Targeting the two types of bibliometric indicators (i.e., terms and bibliographical information), we performed two sets of analyses, respectively, that is, key entity analysis and topic analysis.

A. Key entity analysis

We employed co-occurent statistics between affiliations, between countries and regions, and between research sources to investigate the key entities involved in this global discussion on AI ethics. These entities include 1) universities and research institutions, with their geographical distribution, and 2) journals and their citation patterns.

B. Topic analysis

Topic analysis in this study comprises two parts: 1) profiling the topical landscape of AI ethics in the literature via a topical hierarchical tree [40]; and 2) tracking how concerns about AI ethics have changed over time using a scientific evolutionary pathways (SEP) approach [15].

A topical hierarchical tree (THT) is a network-based algorithm that identifies a hierarchical relationship hidden behind research topics. The assumption is that, in a hierarchical structure, 1) the relationships between a superior and its subordinates are stronger than the relationships with its neighbors, and 2) superiors receive dominant attention compared to their subordinates. When measuring these relationships in this study as the prevalence of terms in a corpus of documents, the THT approach exploits the algorithm for the maximum spanning tree to retrieve the largest undirected graph from a weighted term co-occurrence network. Then, in that graph, we set the terms which receive high prevalence as superiors and their connected terms which receive low prevalence as their subordinates, with directed edges (i.e., arrows) starting from superiors to subordinates. The output of the THT approach is a list of topics with their hierarchical relationships, and we visualized this in the form of mind maps.

The key assumption of the SEP approach is that the accumulative changes of established scientific inventions will trigger scientific evolution once such changes achieve a significant level. Thus, the SEP simulates a corpus of bibliometric records as a bibliometric stream based on their 'publication year' and tracks the change in a topic by monitoring its feature space and the distribution of these features in sequential time slices. When geometrically assuming a topic is a circle with a centroid and a boundary, the SEP measures the Euclidean distance between the centroid of a topic and that of all 'coming' sub-topics generated in a current time slice. When the Euclidean distance exceeds the boundary of the topic, the SEP identifies that sub-topic as a descendant and the original topic is its predecessor. The result of the SEP approach is a list of topics and their predecessor-descendant relationships, and we considered each topic is a node and those relationships as directed edges between them. We used Gephi [41] to visualize the SEPs as a network and its integrated function of community detection 'modularity' [42] groups similar and proximate nodes as a community for further understanding.

Incorporating the results of these topic analyses, we identified a list pairing potential conflicts between current AI techniques and special ethical and privacy issues, which might provide certain insightful knowledge to guide the AI community in future fundamental research and technological development.

4. Results

We investigated AI ethics and privacy by analyzing the articles published in the past few decades (as detailed in our search strategy and data information in Table 1), and we answered the following questions: who are the key players (e.g., research institutions and universities, countries/regions, and research communities) contributing to the research on the ethical and privacy issues relating to AI, from a topical perspective what are AI ethics in detail and how has the interest of academia in these issues changed over time, and what are current emerging issues relating to AI ethics.

4.1. Key players contributing to the work on AI ethics

We utilized the full bibliographical information provided by the WoS Core Collection and analyzed the #3 dataset to identify 1) the key players contributing to the work on the ethical and privacy issues relating to AI and their collaborative patterns and geographical distribution, and 2) disciplinary interactions in relation to the ethical and privacy issues surrounding AI, and the key sources cited by AI ethics-related research articles. Note that #3 is a sub-dataset of #1 but considering the WoS Core Collection database for indexing ‘high-quality’ research articles, despite possible missing data, this key entity analysis should be representative of the standard of current research on AI ethics.

Table 2 lists the top 15 most productive⁹ affiliations (including their countries) contributing to the work on AI ethics and Figures 3 and 4 visualize the collaborative patterns between these affiliations (a total of 3,377 affiliations) and between the top 30 countries/regions, respectively. The following observations can be made: 1) the USA dominates in this area in both the total number of relevant publications and the number of productive affiliations; 2) English-speaking countries, such as the UK, Australia, and Canada have a strong interest in AI ethics and the fact that the top 15 productive affiliations are from English-speaking countries further supports this observation; 3) China is ranked the third most productive country however no other Asian countries appear in this list. Comparably, European countries such as Germany, the Netherlands, Italy, Spain, and France, as a union, produce a large volume of research on this topic.

Table 2. Top 15 affiliations and countries contributing to the work on AI ethics

| | # R | Affiliation | Country | | # R | Country |
|---|-----|----------------|---------|---|------|-------------|
| 1 | 70 | Univ Oxford | UK | 1 | 1011 | USA |
| 2 | 43 | Stanford Univ | USA | 2 | 496 | UK |
| 3 | 37 | Univ Edinburgh | UK | 3 | 266 | China |
| 4 | 30 | MIT | USA | 4 | 216 | Australia |
| 5 | 30 | UCL | USA | 5 | 198 | Canada |
| 6 | 29 | Univ Toronto | Canada | 6 | 193 | Germany |
| 7 | 28 | NYU | USA | 7 | 153 | Netherlands |

⁹ The productivity in this work was decided based on the total number of articles published by a given entity (e.g., affiliations in Table 2 and publication sources in Table 3).

| | | | | | | |
|----|----|-----------------------|-----------|----|-----|-------------|
| 8 | 27 | Harvard Univ | USA | 8 | 136 | Italy |
| 9 | 25 | Univ Penn | USA | 9 | 134 | Spain |
| 10 | 25 | Univ Sydney | Australia | 10 | 106 | France |
| 11 | 24 | McGill Univ | Canada | 11 | 90 | Switzerland |
| 12 | 24 | Nanyang Technol Univ | Singapore | 12 | 80 | India |
| 13 | 24 | Natl Univ Singapore | Singapore | 13 | 77 | Russia |
| 14 | 24 | Univ Michigan | USA | 14 | 60 | Denmark |
| 15 | 23 | Univ British Columbia | Canada | 15 | 55 | Belgium |

Figure 3 provides a bird's eye view revealing the collaborative patterns of all 3,377 affiliations contributing to the work on AI ethics. In line with our observations from Table 2. The universities from English-speaking countries, particularly the USA and UK, are at the center of the map, which indicates they are leading these collaborative networks, but the European universities (e.g., KU Leuven, Tech Univ Munich, Univ Porto, and Charite Univ Med Berlin) concentrate on their relatively small groups and China (Chinese Acad Sci, Peking Univ, and Baidu) is also located at a marginal area of the co-authorship network. The active role played by the leading universities (e.g., MIT, Univ Penn, Univ Oxford, Natl Univ Singapore, and Univ Edinburgh) in conducting research on AI ethics may indicate the increasing interest in this field from academia and its urgency.

Figure 4 shows the co-authorship map for the top 30 countries and regions in relation to the research on AI ethics. The USA produces the most research on discussing AI ethical issues and its collaborative network covers almost all the countries/regions in this map and it has particularly strong ties with the UK, Canada, Australia, China, Germany, and the Netherlands. However, while domestic collaboration is obviously the key pattern in the leading countries (e.g., approximately 66% in the USA, 52% in the UK, and 62% in China), we also observe several European countries have a preference for international collaboration, such as Austria, Belgium, and Sweden – the proportion of their international collaboration achieves almost 60%.

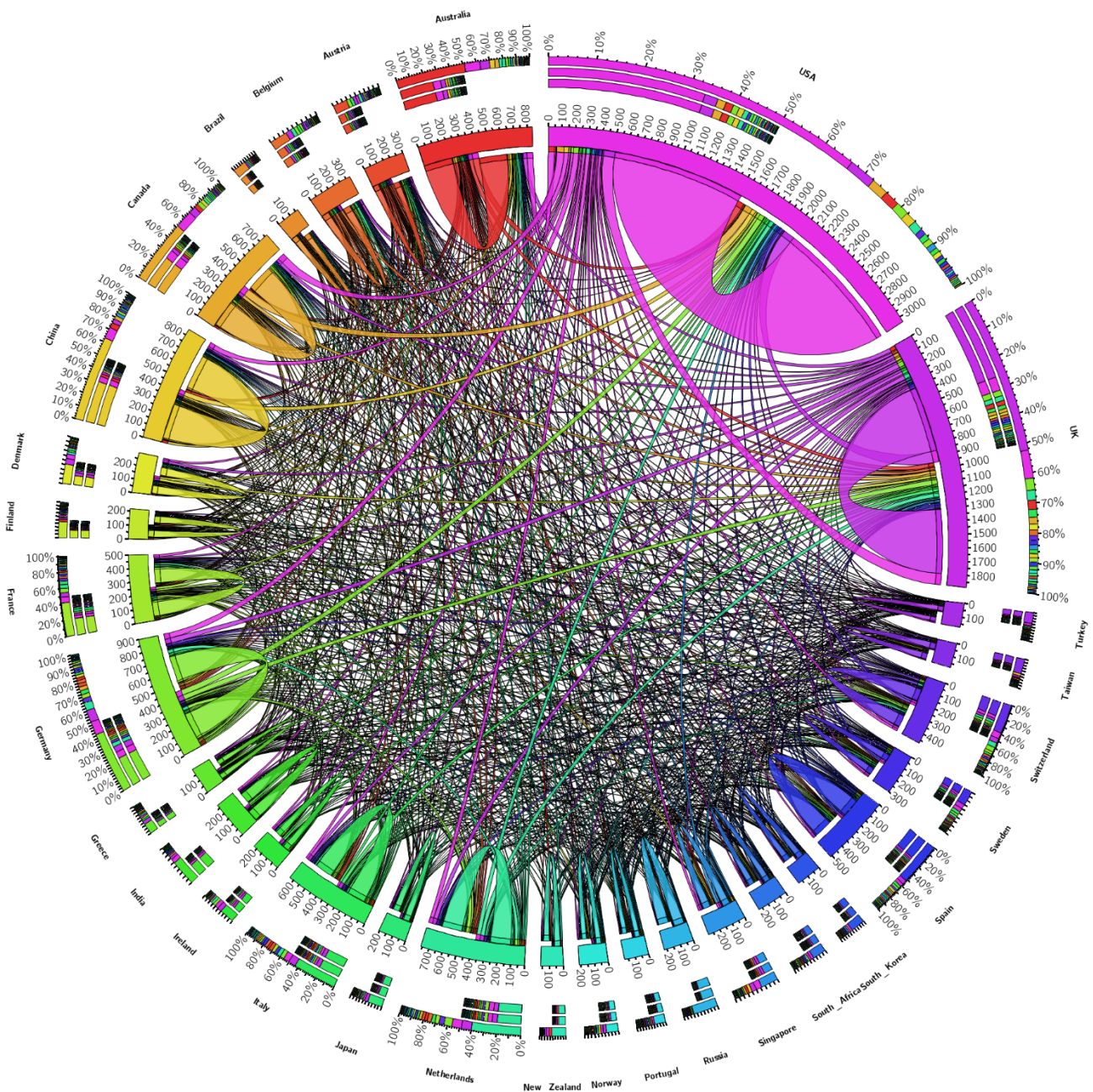


Figure 4. Co-authorship map for key countries and regions in relation to the research on AI ethics (visualization tool: Circos [44])

Note that in a map generated by Circos in this paper (i.e., Figures 4 and 9), one slice with a unique color represents one entity (e.g., countries/regions, and topics), and the ribbon link between two slices indicates the strength of their co-occurrence. Specifically, the self-linked ribbons in Figure 4 represent domestic collaborations within a country/region.

Table 3 shows the publication sources (e.g., research journals and magazines) which publish research on AI ethics, as well as the interactions between the WoS subject categories of these sources. The 3,259 articles in #3 dataset were published in 1,936 publication sources, including journals, conference proceedings, and magazines, and Table 3 lists the top 24 most productive publication sources on AI ethics. Except for three conference proceedings and one magazine, most of these publications are in research journals and are from diverse disciplines, such as computer science, medical science, biology, and media. As reflected in the names of these publications, one common interest of these journals is to investigate the societal impact (e.g., ethics, law, crimes, and sustainability) of science and technology.

Table 3. Publication sources with more than 10 articles on AI ethics

| #R | Publication Source | #R | Publication Source |
|----|--|----|---|
| 1 | 56 AI & Society | 13 | 13 Proceedings of the 2019 AIES |
| 2 | 41 Big Data & Society | 14 | 13 BMJ Open |
| 3 | 40 Science and Engineering Ethics | 15 | 12 AI Magazine |
| 4 | 34 Ethics and Information Technology | 16 | 12 Journal of Information Communication & Ethics in Society |
| 5 | 23 Computer Law & Security Review | 17 | 12 Russian Journal of Criminology |
| 6 | 23 Journal of Medical Internet Research | 18 | 11 Asian Bioethics Review |
| 7 | 23 Minds and Machines | 19 | 11 OMICS |
| 8 | 19 IEEE Access | 20 | 11 Proceedings of the 2019 ECIAIR |
| 9 | 18 Proceedings of the 2018 AIES | 21 | 11 Sustainability |
| 10 | 17 Philosophical Transactions of The Royal Society A | 22 | 10 Journal of Bioethical Inquiry |
| 11 | 16 BMC Medical Ethics | 23 | 10 New Media & Society |
| 12 | 15 Information Communication & Society | 24 | 10 Social Media + Society |

Every journal covered by the Web of Science Core Collection is assigned to at least one of 254 subject categories. We retrieved 199 WoS categories from these 1,936 publications, revealing a multi-disciplinary interest in AI ethics, and we visualized their co-occurrence relationships in Figure 5. We summarize and discuss the following key observations:

- The three WoS categories (“computer science & artificial intelligence”, “computer science, theory & methods”, and “computer science & information systems”),

which build the core knowledge pillars on the fundamental research and applications of AI, together with “engineering, electrical & electronic”, “computer science, hardware & software”, and “computer science, software engineering”, form the technical backbone of AI (red nodes). Its key application areas in “medical informatics” and “health care sciences & services” further extend this technical scope (light green nodes).

- Ethical issues (purple and grey-blue nodes) are discussed in extensive categories of social sciences, such as “ethics”, “history & philosophy of science”, “philosophy”, “medical ethics”, and “social sciences, biomedical”. As supplementary sources, “information science & library science”, “management”, and “economics” provide analytic approaches (brown nodes), while the engagement of “regional & urban planning”, “environmental studies”, “political science”, and “education & educational research” involves new application scenarios (blue nodes).

To track the knowledge flow through the citation behaviors of the 3,259 articles, we collected their references and retrieved 51,431 journals. The co-occurrence relationships of these cited journals are visualized in Figure 6, providing a new angle to identify the research communities contributing to the research on AI ethical issues. We raise the following points:

- Relatively clear boundaries among five communities indicate an established knowledge system on AI ethics, namely computer science (purple nodes), information systems and management (red nodes), medical sciences and multi-disciplinary studies (green nodes), and law and general magazines (blue nodes).
- Leading journals play an active role in bridging cross-disciplines, and the publication of AI ethics in reputable newspapers and magazines assists in increasing the awareness of the general public in these issues. The following publications construct the backbone of this knowledge system: *Nature*, *Science*, *PNAS*, *PloS One*, *JAMA*, *New England Journal of Medicine*, *Lecture Notes in Computer Science*, *Communication of ACM*, *Big Data & Society*, *Information Communication & Society*, *Guardian*, *New York Times*, *Ethics and Information Technology*, and *Science and Engineering Ethics*.

In conclusion, in this section we identified the key players (e.g., research institutions and communities) which contribute to the research on the ethical issues surrounding AI and the countries/regions where this research is being undertaken. Such insights draw a landscape to support the understanding of “who” has been involved in the study of AI ethics and “how” they have contributed to this topic. In particular, we highlighted the role of cross-disciplinary research publications (e.g., *Communications of ACM*, and *Lecture Notes in Computer Science*), multi-disciplinary research journals (e.g., *Nature* and *Science*), and newspapers (e.g., *New York Times*) in gradually transferring technical AI knowledge to inform public concerns on ethics.

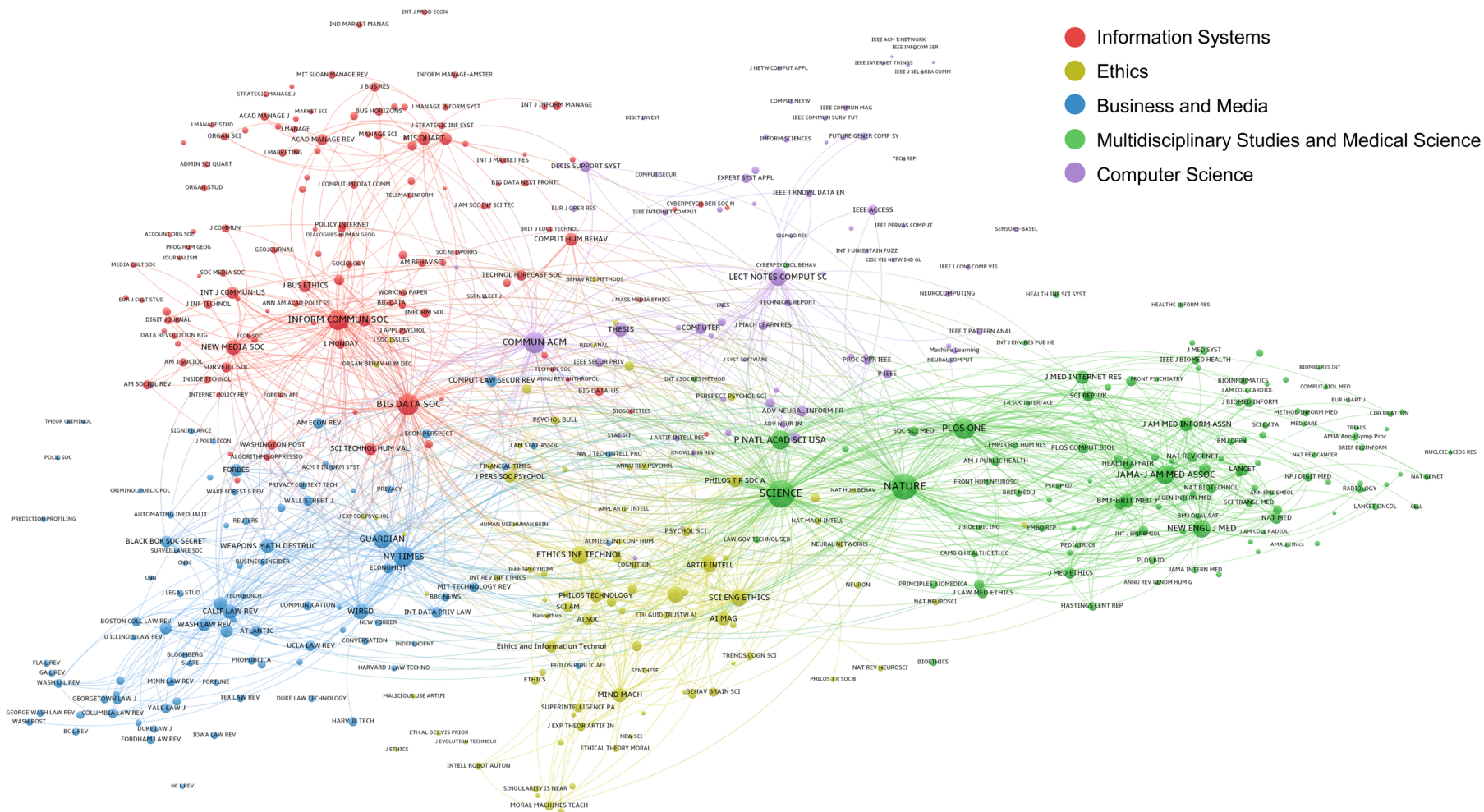


Figure 6. Co-citation network for journals cited by research articles on AI ethics (visualization tool: VOSViewer [43])

4.2. Landscapes and evolution of AI's ethical topics

In this section, we move our foci to topic analyses by analyzing #1 dataset collected from the WoS All Databases. We initially retrieved 93,364 terms from the combined titles and abstracts of the 4,375 articles, and we conducted a term clumping process [13] to remove noise and consolidate the technical synonyms, reducing the total number of terms to 52,054. Then, we used the 2,163 terms appearing in more than 2 articles as the core set of terms to generate the topical hierarchical tree (THT) shown in Figure 7 and the map of the scientific evolutionary pathways (SEP) shown in Figure 8.

Figure 7 enhances the understanding of the details of AI ethical issues, especially the connections between specific AI techniques and ethical concerns. Among its 71 nodes, the THT lists 27 AI techniques (e.g., machine learning) and AI-driven applications, devices, and products (e.g., robots and autonomous vehicles), 28 ethical topics (e.g., fairness and discrimination), and 16 societal topics (most of them in relation to medical and healthcare issues). The four main branches of this THT represent four major issues relating to AI ethics, that is, #1 AI techniques and potential ethical issues, #2 technological and political implications of AI ethics, #3 data privacy, and #4 privacy in healthcare. We discuss these four issues in detail:

#1 AI techniques and potential ethical issues: Figure 7 reveals the key AI techniques that may raise ethical concerns, such as machine learning (including deep learning, computer vision, neural networks, natural language processing, etc.), ontologies, communication technologies, and neuroscience¹⁰. Machine learning, one of the key areas in AI, shares close connections with almost all AI techniques, and thus attracts the most attention in this THT and are connected with all ethical issues, such as fairness, discrimination, liability, frauds, and criminals¹¹. It is easy to explain these cases. For example, applying AI models to make decisions entails justifiable “right to a well-calibrated machine decision” [45, 46], AI-driven fraud in social media, political elections, and financial markets (e.g., fake videos and identifications manipulated by AI techniques, such as image processing and face recognition) have become a major concern [47]. How to validate AI recommendations with human knowledge in actual cases, such as clinical practice, is challenging both the AI community and the receptivity of the general public [48]. A brand-new topic, brain computer interface is attracting increasing attention from the public, and ethical issues (such as privacy) and related regulations are appearing in public reading materials [49].

#2 technological and political implications of AI ethics: As an extension of the ethical issues in #1, #2 further extends AI's influence from ethics to the broad society through specific technological and political implications, such as sustainability, responsibility, and digitalization. From the perspective of a complex ecosystem, these societal

¹⁰ Neuroscience here mostly refers to techniques of brain computer interface.

¹¹ We note that fairness is one constraint in evaluating reinforcement learning approaches and fraud detection is a specific task of machine learning, and thus these variations might introduce noise to our analysis.

reactions could be the resilient progress of an ecosystem responding to disruptions introduced by AI techniques and their resulting ethical issues [50].

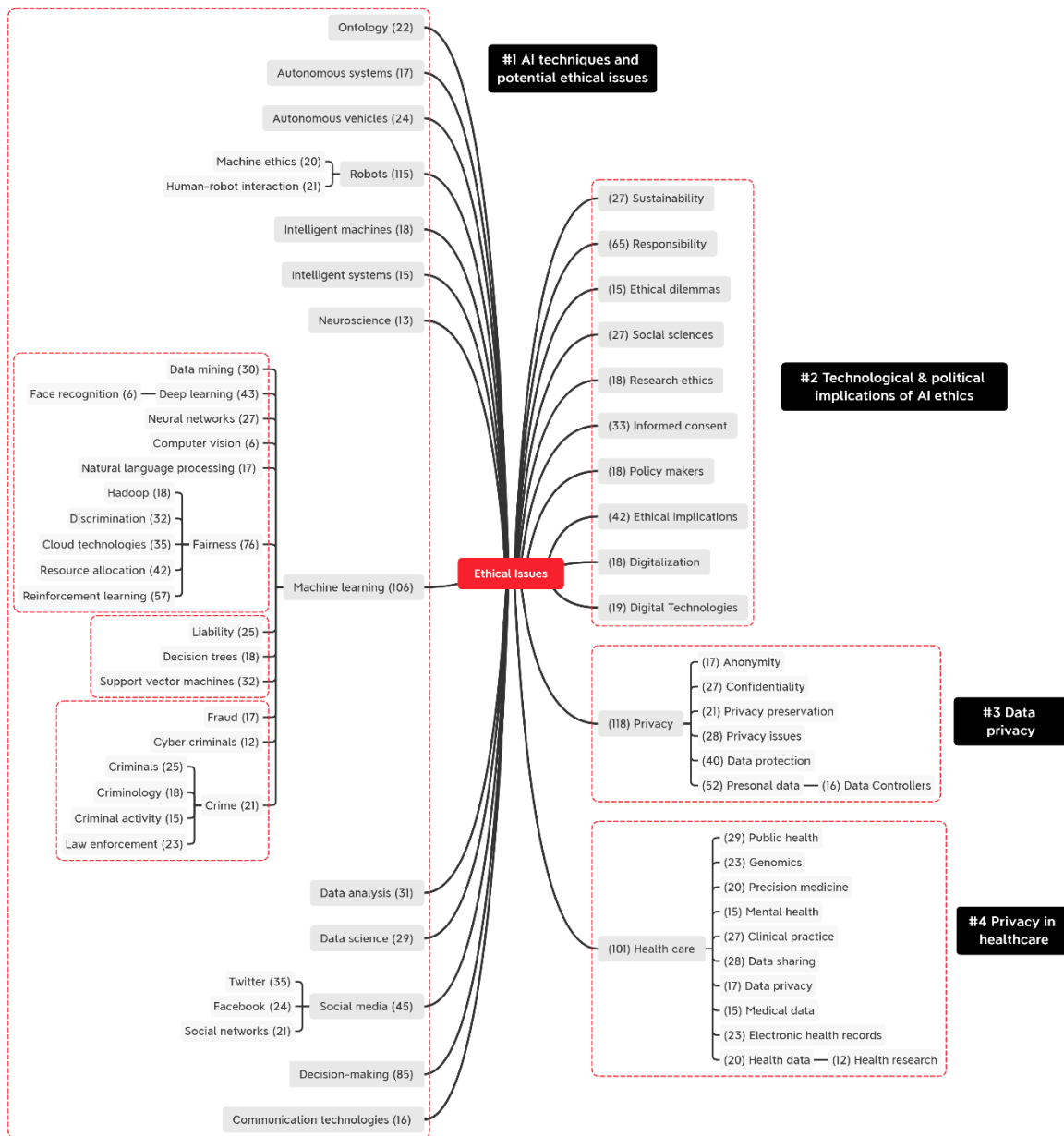


Figure 7. Topical hierarchical tree on AI ethics

Note that the number in the brackets after each topic represents the importance of the topic, measured by the frequency of the co-occurrence between the topic and its upper-level topic.

#3 data privacy and #4 privacy in healthcare: #3 and #4 are a specific case of AI ethics. The big data boom initially activated the public's concerns on data privacy, where the illegal exposure of personal data, particularly those linked with social media, occurred, e.g., the Facebook case in Footnote 1. Furthermore, while analyzing health

data (e.g., electronic health records), including clinical trials and gene sequencing data provides evidence for precision medicine, privacy concerns in medical and healthcare sectors then become not only a societal issue but are also a threat to national strategies and the sustainability and balance of nature [51].

To further explore the details of these AI ethical issues and their evolutionary relationships over the past few decades, the topical evolutionary pathways on AI ethics are visualized in Figure 8. We set ‘expert systems’ as the starting point of the evolutionary pathways, considering it is a representative AI technique/application in the 1990s and before. Seven communities, represented by different colors in Figure 8, uncover diverse interests and emphases in AI techniques, applications, and related ethical concerns. They are #1 expert systems (dusty yellow), #2 criminal investigation (macaron blue), #3 machine ethics (grass green), #4 anonymity (light purple), #5 decision making (ocean blue), #6 health care (orange), and #7 clinical practice (peach red). We observed certain findings and discussed these as follows:



Figure 8. Topical evolutionary pathways on AI ethics (1977-2020)

Expert systems (#1) represent the interactions between AI techniques (and information technologies in the early years) and human knowledge, and thus, together with practical cases such as energy efficiency, it seems that fairness issues mainly appear in this path.

Criminals (#2) involves criminal justice, crime analysis, cyber criminals, and liability. This relatively new community started in 2014, and its two large branches appeared

in 2016 and after. One interesting aspect here is the involvement of face recognition techniques in cyber criminals, and the ‘deepfake’ story ¹² may well endorse this observation, in which an AI mobile app can insert faces in place of film and TV characters and may result in fraud by defeating the ‘Face ID’ function in smart phones. The other aspect for computer vision is its use in law enforcement (e.g., surveillance systems) for crime detection in national security activities. However, such techniques violate personal privacy in these practical uses.

The study of machine ethics (#3) results in a timeline showing how public concerns about social media privacy have changed over time – e.g., from illicit activities of social media platforms in 2016 to responsibility one year later, and from a governance framework in 2019 to regulations in relation to ethical behaviors and dimensions in 2020.

From anonymity in 2004, #4 develops into a relatively broad scope of ethical concerns in research and health data (e.g., potential influence of electronic health records), data protection, and privacy. In the other main branch of this community, from a technical perspective, autonomous vehicles and cognitive capabilities could act as open data sources and benefits, but interestingly, how to protect sensitive information in open data initiatives has become an issue as well and cybersecurity further strengthens such protection [52].

#5 is a community investigating the traditional base of information systems, in which multi-agent systems and intelligent systems were involved before 2014. After this, increased constraints such as accountability, confidentiality, and sustainability to evaluate the capabilities of information systems indicate the emerging interests of this research community. Particularly, rooted in accountability, new concerns on monetization of data were raised in 2019, inspiring global-wide debates on diverse aspects, from political governance to legal and financial regulations.

#6 is an extension of decision-making in diverse scenarios such as health care and medical data and with diverse theories, concepts, and techniques, such as ontology, neuroscience, and game theory, but in this path, human morality, together with human factors and misleading information, is specifically highlighted. Another highlight here regards to neurosciences. As we discussed in Figure 7, brain computer interface may align with this topic, which analyzes brain signals and makes decisions for human beings, and such activities attract comments on human morality – [49] quotes from one of its interviewed ethicists, a device of brain computer interface “was more than a device...the company owned the existence of this new person”. Thus, it is critical to discuss how to regulate these new AI devices.

#7 contains the largest number of emerging topics generated in recent years and ethical issues in clinical practice are a key concern not only to academia but also to the general public. Like our discussion of #4, such concerns mostly revolve around the illegal use of various personal data, such as health records, clinical data, and genomic

¹² See details of this news on the website: <https://www.bbc.com/news/technology-49570418>

data, as well as data sharing and security. In 2020, following the topic of bioethics, gene editing, the winner of the 2020 Nobel Biology prize, attracted the attention of this community.

As a conclusion for the SEP, 1) data privacy is an urgent topic relating to AI ethics, particularly when data contain sensitive personal information, with clinical trials and genomics; 2) the increasing threats and fears in relation to AI-driven fraud and cybercrime are drawing attention; and 3) the reliability, transparency, and fairness of AI models are still unsolved issues.

As discussed and highlighted in the THT and SEP, of particular interest to the AI community is the discovery of potential conflicts between AI techniques and specific ethical issues, and thus, referring to our search strategy (Table 1) and terms appearing in Figures 7 and 8, we selected 15 AI technique-related terms and 17 AI ethics-related terms, and visualized their co-occurrence relationships in Figure 9. We discuss these AI techniques and their closely connected ethical issues in the following:

- Machine learning as a representative technique, including deep learning, reinforcement learning, and neural networks, touches all 17 ethical issues, particularly, fairness, accountability, and privacy. Despite different emphases, data mining and cloud technologies follow similar patterns. In this area, all concerns come to the balance between AI decisions and the mechanism behind that decision (e.g., data collection and algorithmic transparency).
- Computer vision (including face recognition and image processing techniques) is raising concerns from the general public. These are directly linked with crime (regarding manipulated fake images and videos and surveillance systems used for national and domestic security detection) and accountability issues.
- Robots, as an engineering-driven AI application, draw attention in relation to machine ethics, responsibility, accountability, liability, and privacy, as do autonomous vehicles. Since political regulations for those intelligent machines lag its technological progress, the general public worries about the reliability of these new technologies (e.g., the safety of an autonomous car), and broad ethical and moral issues (e.g., how shall we charge a machine with a crime, and who is liable for a failure).
- Blockchain techniques, an interdisciplinary area with both AI hardware and software, attract criticism in relation to accountability and sustainability. In fact, from a public point of view, blockchain techniques are heavily involved in the internet of things, and thus, compared to traditional ethical issues, sustainability is a special concern in this area.
- Neuroscience, as a discipline for techniques in brain computer interface, represents current AI activities in collecting personal information, in clinical trials, healthcare records, genomic data, and brain signals. Despite great potential in benefiting human beings in precision medicine, disability and accessibility services, and smart home, critical concerns align with privacy and responsibility.

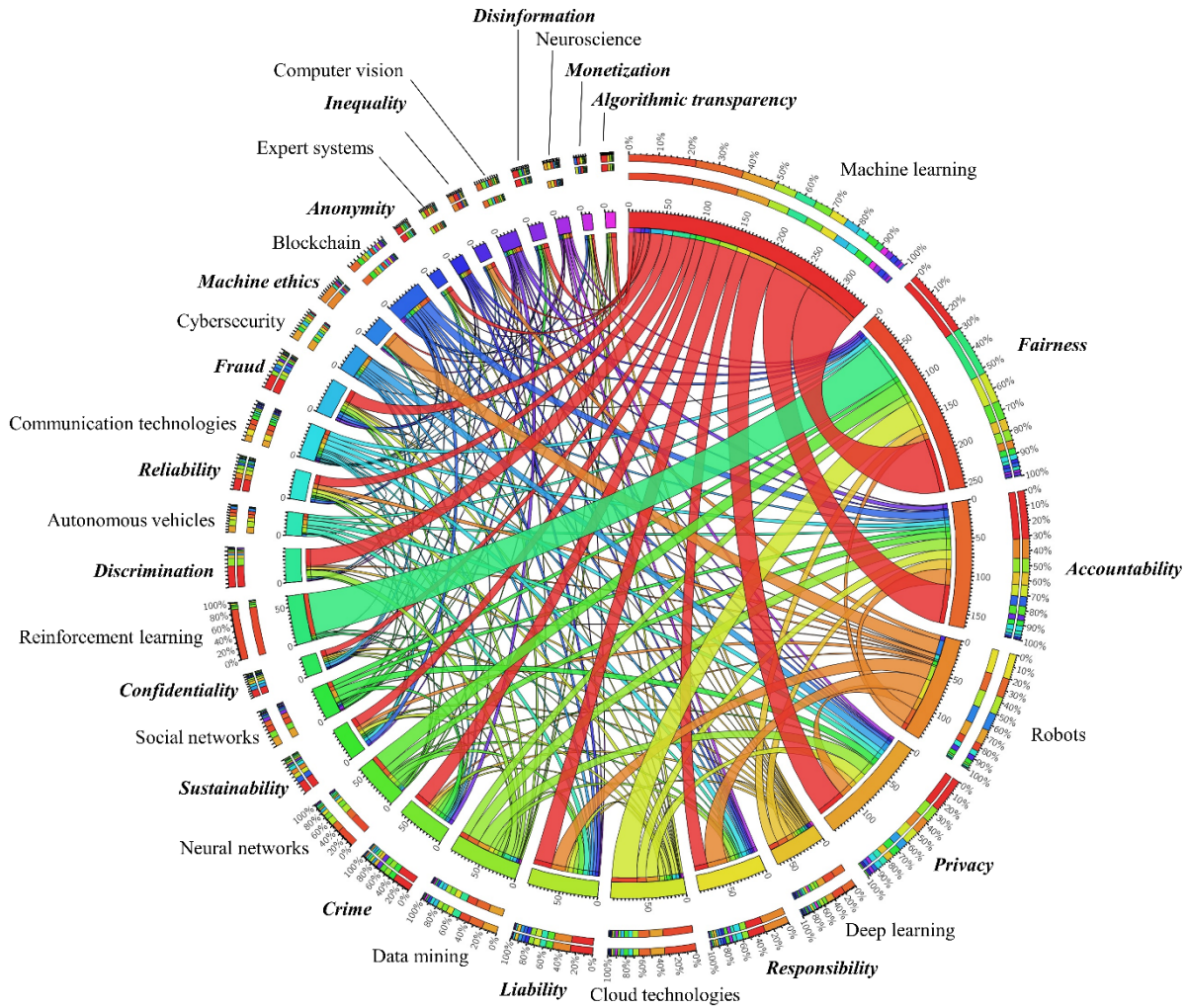


Figure 9. Co-occurrence map between 15 key AI techniques and 17 ethical topics (visualization tool: Circos [44])

Note that bold and italic labels represent ethical topics and other labels present topics of AI techniques.

4.3. Current emerging issues in AI ethics and privacy

The specific interest of digging out emerging issues in AI ethics leads us to timely articles published in the three leading outlets (i.e., *Nature*, *Science*, and *PNAS*), and thus, we analyzed the 53 articles in dataset #2 but we removed articles published before 2015, then manually read the remaining 46 articles and selected 34 articles which directly touch on AI ethical issues. Interestingly, most of these articles are opinions, news, and comments, and *Nature* is the key publication (25 articles). Compared to research articles, these “informal” types of articles might reflect the increasing interest of the public to AI ethics, and such opinions and comments could be some rapid re-actions to emerging ethical issues in relation to AI (but may need further extensions and studies to enrich them to full research articles). Given this circumstance, we consider this section as a complementary study of Section 4.2, and the main purpose here is to explore current emerging issues in AI ethics.

With the involvement of manual intervention, we categorized the 34 articles into the following main topics in Figure 10.

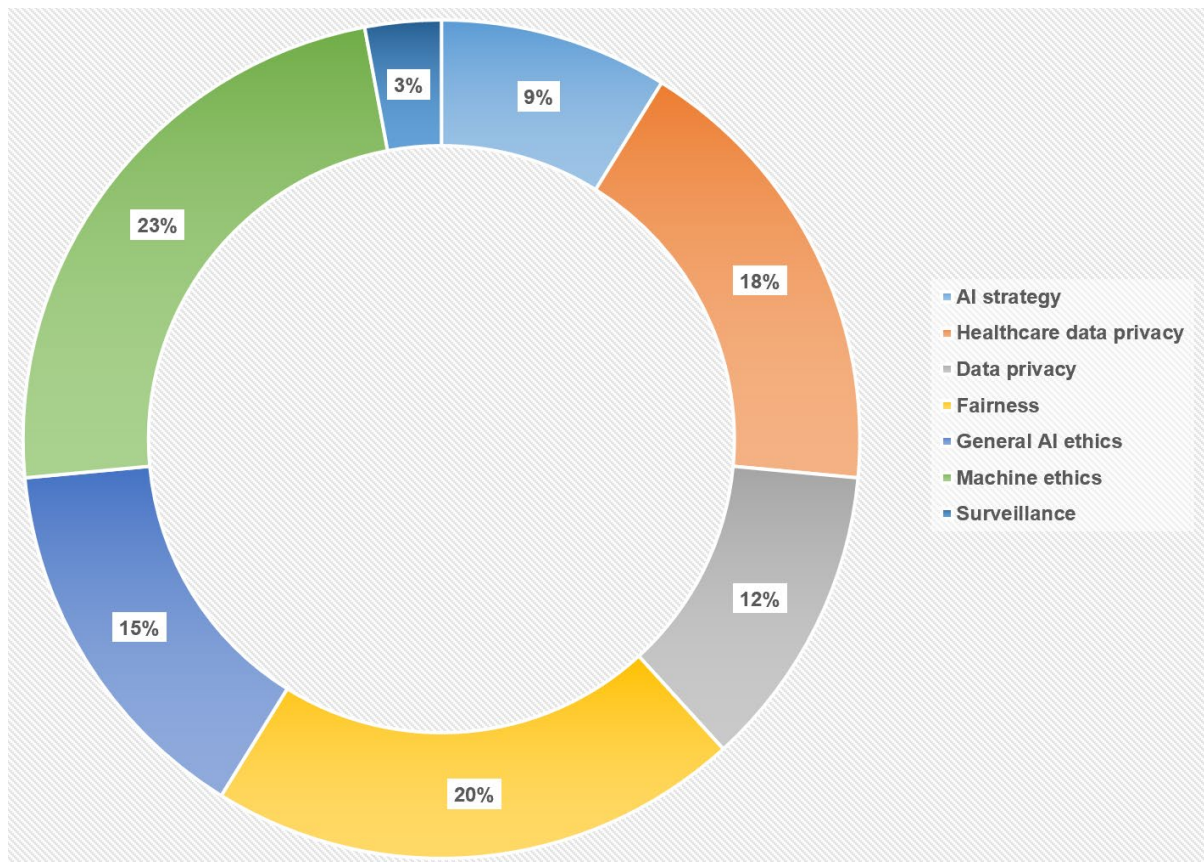


Figure 10. Current urgent topics in relation to AI ethics

Privacy issues (30%) are one of the key emerging concerns. This is consistent with the position of the UNESCO in its recent draft of the *Recommendation on the Ethics of Artificial Intelligence*, which, as mentioned in Section 2.1, listed privacy as one of key principles of AI ethics. Specifically, concerns were expressed in 18% of the articles about healthcare data privacy, with a focus on issues such as the balance of governance on public health control and data privacy for patient records, disease monitoring, and genomic data. The other 12% of articles expressed concerns relating to big data privacy. One specific interest comes to the observation here, which reveal that healthcare data privacy constitutes a separate topic from topic data privacy. In fact, this result is consistent with the current privacy governance and regulatory structure in Australia. Using the privacy laws in State of New South Wales (NSW) as an example, there are two major statutory laws governing privacy and personal information protection in NSW. One is the Privacy and Personal Information Protection Act 1998 (NSW) (hereinafter 'PPIPA'). The other is the Health Records and Information Privacy Act 2002 (NSW) (hereinafter 'HRIP'). The PPIPA offers protection to all personal information except health information. S4A of PPIPA explicitly excludes the "health information" under the HRIP from definition of personal information under

the PPIPA¹³. In contrast, the HRIP focuses on health information with the purpose of “promoting fair and responsible handling of health information” in particular¹⁴. Such a governance structure (personal information + health information) is arguably consistent with the structure of our observations in the series of topic analyses (data privacy + healthcare data privacy). This may arguably serve as *prima facie* evidence on the accuracy and reliability of our system.

Other concerns are mainly related to machine ethics (23%) and fairness (20%). Specifically, machine ethics touches on a wide range of topics relating to the morality of intelligent machines (e.g., AI cars), how to uphold human rights with robots, and the consciousness of machines. These discussions reflect the potential fears of the general public relating to these unknown but extremely smart machines and the dilemma between technology and ethics. On the other hand, fairness indicates the unease of the general public as to whether AI models can generate fair decisions in diverse scenarios. Articles related to AI strategy mainly talk about how to regulate this new AI world in a power-shifting theme (e.g., how to seek a balance between human beings and intelligent machines) and how shall national strategies and military actions involve in the development of AI techniques.

In addition to a general discussion on the issues surrounding AI ethics, surveillance seems to be an increasing concern, where the authors of these articles call for the review and regulation of AI surveillance systems, regardless of whether they are used for national security, industrial monitoring, or research/individual use.

5. Conclusions

In this paper, we analyzed articles indexed in the Web of Science to investigate the ethical issues surrounding AI which are becoming an increasing concern not only to academia but also the general public. We identified the key affiliations, countries/regions, and research communities which contribute to the research on the ethical issues surrounding AI via a series of co-occurrence analysis on bibliographical information of the collected articles. We then profiled the AI ethical issues in both hierarchical and time dimensions via intelligent bibliometric approaches, including topical hierarchical trees and scientific evolutionary pathways, which helped us answer the questions as to what are the specific ethical issues of concern relating to AI are and how has society’s interests in these issues changed over time. We specifically concentrated on the most recent articles published in *Nature*, *Science*, and *PNAS*, and discovered the current urgent issues of interest to the AI ethics community. We expect that the insightful findings identified in this study will support the understanding of the AI community in relation to AI ethics, especially in profiling the hidden ethical issues behind specific AI techniques. Such findings, bridging the knowledge base of AI

¹³ S4A states: ‘Except as provided by this Act or the Health Records and Information Privacy Act 2002, the definition of “personal information” in section 4 does not include health information within the meaning of the Health Records and Information Privacy Act 2002’.

¹⁴ See s3 of the HRIP.

techniques and ethical discussion in public debates, will be of interest to audiences in science policy, technology management, and public administration.

5.1. Key findings

Referring to Tables 2 and 3, this paper found that the key contributors to the research on the ethical issues relating to AI were English-speaking countries such as the USA, UK, Australia, and Canada. In comparison, China and the European countries contribute to this research area as well, but their key research institutions are not as equally appealing as those of English-speaking countries. According to Figure 4, intriguingly, those countries making the major contribution to the research on AI ethical issues, namely the USA, UK, and China, mostly engage in domestic collaboration, however certain European countries, such as Austria, Belgium and Sweden, seem to prefer international collaboration.

In Figures 5 and 6, the ethical issues relating to AI cover a wide range of disciplines (i.e., 199 of the 254 WoS categories), and four research communities play an active role in the research associated with AI ethical issues, namely computer science, business and management, medical science, and law. The involvement of newspapers and magazines in publishing research on AI ethical issues indicates the interest of the general public in these matters.

In terms of topic analysis in Sections 4.2 and 4.3, key AI techniques such as machine learning, data analysis, robots and intelligent systems, and cloud technologies, generate concerns about the ethical issues relating to AI. Fairness, as well as discrimination, are among those key concerns because AI models are applied in decision support in diverse scenarios. Data privacy, particularly in the healthcare and medical sectors, is a cause of increasing concern. Cybercrime and fraudulent behavior are particularly concerning in the absence of appropriate support from the law and regulations. Machine ethics are mostly related to robots, autonomous cars, and intelligent machines, highlighting a balance between machine consciousness and human rights.

5.2. Limitations and future directions

We anticipate the future directions of this study by addressing the following limitations: 1) Instead of research publications, further broad discussion on AI ethics might take place on social media platforms, such as Twitter, Facebook, and mainstream medias (e.g., BBC and CNN), and thus, future studies could integrate these sources with traditional bibliometric data broadening the study for more comprehensive results; 2) Although the WoS All Databases provide a rich collection of various types of articles, not all articles (e.g., news) contain abstracts and missing abstracts might influence the algorithmic precision and topical coverage of our topic analyses; 3) The labeling strategy (e.g., how to label a node or a tree branch) is still a challenging issue in text analysis-based bibliometrics, and the balance between algorithmic logic and semantics may heavily influence the understanding of the results; and 4) since the key idea of bibliometrics is to quantitatively explore items, patterns,

and relationships from scientific documents, answering questions of “why”, “how”, and “so what” may require interactive and extensive engagements with experts in AI ethics. This could be one of our long-term collaborations combining researchers from the AI community and the legal community to uncover the hidden mechanisms and reasons behinds those explorative results.

Acknowledgements

This work is supported by the Australian Research Council under Discovery Early Career Researcher Award DE190100994.

References

- [1] R. Harper, *Inside the smart home*. Springer Science & Business Media, 2006.
- [2] A. Walter, R. Finger, R. Huber, and N. Buchmann, "Opinion: Smart farming is key to developing sustainable agriculture," *Proceedings of the National Academy of Sciences*, vol. 114, no. 24, pp. 6148-6150, 2017.
- [3] F. S. Collins and H. Varmus, "A new initiative on precision medicine," *New England journal of medicine*, vol. 372, no. 9, pp. 793-795, 2015.
- [4] M. S. Hossain, G. Muhammad, and N. Guizani, "Explainable AI and mass surveillance system-based healthcare framework to combat COVID-19 like pandemics," *IEEE Network*, vol. 34, no. 4, pp. 126-132, 2020.
- [5] J. Bossmann. "Top 9 ethical issues in artificial intelligence." World Economic Forum. (accessed October 26, 2020).
- [6] V. C. Müller, "Ethics of artificial intelligence and robotics," 2020.
- [7] Y. Zhang, G. Zhang, H. Chen, A. L. Porter, D. Zhu, and J. Lu, "Topic analysis and forecasting for science, technology and innovation: Methodology and a case study focusing on big data research," *Technological Forecasting and Social Change*, vol. 105, pp. 179-191, 2016.
- [8] Y. Zhang *et al.*, "Does deep learning help topic extraction? A kernel k-means clustering method with word embedding," *Journal of Informetrics*, vol. 12, no. 4, pp. 1099-1117, 2018.
- [9] M. Wu, Y. Zhang, G. Zhang, and J. Lu, "Exploring the genetic basis of diseases through a heterogeneous bibliometric network: A methodology and case study " *Technological Forecasting and Social Change*, 2020.
- [10] Y. Zhang, M. Wu, Z. Hu, R. Ward, X. Zhang, and A. Porter, "Profiling and predicting the problem-solving patterns in China's research systems: A methodology of intelligent bibliometrics and empirical insights," *Quantitative Science Studies*, 2020.
- [11] D. Cetindamar, T. Lammers, and Y. Zhang, "Exploring the knowledge spillovers of a technology in an entrepreneurial ecosystem—The case of artificial intelligence in Sydney," *Thunderbird International Business Review*, vol. 62, no. 5, pp. 457-474, 2020.
- [12] Y. Zhang, X. Zhou, A. L. Porter, and J. M. V. Gomila, "How to combine term clumping and technology roadmapping for newly emerging science & technology competitive intelligence: “Problem & Solution” pattern based semantic TRIZ tool and case study," *Scientometrics*, vol. 101, no. 2, pp. 1375-1389, 2014.
- [13] Y. Zhang, A. L. Porter, Z. Hu, Y. Guo, and N. C. Newman, ""Term clumping" for technical intelligence: A case study on dye-sensitized solar cells," *Technological Forecasting and Social Change*, vol. 85, pp. 26-39, 2014.

- [14] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," *Advances in Neural Information Processing Systems*, pp. 3111-3119, 2013.
- [15] Y. Zhang, G. Zhang, D. Zhu, and J. Lu, "Scientific evolutionary pathways: Identifying and visualizing relationships for scientific topics," *Journal of the Association for Information Science and Technology*, vol. 68, no. 8, pp. 1925-1939, 2017.
- [16] R. Attfield, *Ethics: an overview*. Bloomsbury Publishing, 2012.
- [17] K. Grace, J. Salvatier, A. Dafoe, B. Zhang, and O. Evans, "When will AI exceed human performance? Evidence from AI experts," *Journal of Artificial Intelligence Research*, vol. 62, pp. 729-754, 2018.
- [18] J.-F. Bonnefon, A. Shariff, and I. Rahwan, "The social dilemma of autonomous vehicles," *Science*, vol. 352, no. 6293, pp. 1573-1576, 2016.
- [19] W. Wallach and C. Allen, *Moral machines: Teaching robots right from wrong*. Oxford University Press, 2008.
- [20] L. Floridi, "Artificial intelligence, deepfakes and a future of ectypes," *Philosophy & Technology*, vol. 31, no. 3, pp. 317-321, 2018.
- [21] C. Cath, S. Wachter, B. Mittelstadt, M. Taddeo, and L. Floridi, "Artificial intelligence and the 'good society': the US, EU, and UK approach," *Science and engineering ethics*, vol. 24, no. 2, pp. 505-528, 2018.
- [22] G. Y. Tian, "Current Issues of Cross-Border Personal Data Protection in the Context of Cloud Computing and Trans-Pacific Partnership Agreement: Join or Withdraw," *Wis. Int'l LJ*, vol. 34, p. 367, 2016.
- [23] C. Kerry, *Protecting Privacy in an AI-Driven World*. Brookings, 2020.
- [24] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," *Nature Machine Intelligence*, vol. 1, no. 9, pp. 389-399, 2019.
- [25] B. Mittelstadt, "AI Ethics—Too principled to fail," *arXiv preprint arXiv:1906.06668*, 2019.
- [26] T. Hagendorff, "The ethics of AI ethics: An evaluation of guidelines," *Minds and Machines*, pp. 1-22, 2020.
- [27] D. Price, "Little Science, Big Science," ed. NY: Columbia University Press, 1963.
- [28] A. Pritchard, "Statistical bibliography or bibliometrics," *Journal of Documentation*, vol. 25, no. 4, pp. 348-349, 1969.
- [29] Y. Zhang, Y. Guo, X. Wang, D. Zhu, and A. L. Porter, "A hybrid visualisation model for technology roadmapping: Bibliometrics, qualitative methodology and empirical study," *Technology Analysis & Strategic Management*, vol. 25, no. 6, pp. 707-724, 2013.
- [30] Y. Zhang, A. L. Porter, S. W. Cunningham, D. Chiavetta, and N. Newman, "Parallel or intersecting lines? Intelligent bibliometrics for investigating the involvement of data science in policy analysis," *IEEE Transactions on Engineering Management*, vol. to appear, 2020.
- [31] D. M. Blei, "Probabilistic topic models," *Communications of the ACM*, vol. 55, no. 4, pp. 77-84, 2012.
- [32] T. Velden, K. W. Boyack, J. Gläser, R. Koopman, A. Scharnhorst, and S. Wang, "Comparison of topic extraction approaches and their results," *Scientometrics*, vol. 111, no. 2, pp. 1169-1221, 2017, doi: 10.1007/s11192-017-2306-1.

- [33] Y. Zhang, H. Chen, J. Lu, and G. Zhang, "Detecting and predicting the topic change of Knowledge-based Systems: A topic-based bibliometric analysis from 1991 to 2016," *Knowledge-Based Systems*, vol. 133, pp. 255-268, 2017.
- [34] J. Guo, X. Wang, Q. Li, and D. Zhu, "Subject–action–object-based morphology analysis for determining the direction of technological change," *Technological Forecasting and Social Change*, vol. 105, pp. 27-40, 2016.
- [35] L. Huang, Y. Zhu, Y. Zhang, X. Zhou, and X. Jia, "A Link Prediction-Based Method for Identifying Potential Cooperation Partners: A Case Study on Four Journals of Informetrics," in *2018 Portland International Conference on Management of Engineering and Technology (PICMET)*, 2018: IEEE, pp. 1-6.
- [36] E. Yan and R. Guns, "Predicting and recommending collaborations: An author-, institution-, and country-level analysis," *Journal of Informetrics*, vol. 8, no. 2, pp. 295-309, 2014.
- [37] Y. Zhang, X. Wang, L. Huang, G. Zhang, and J. Lu, "Predicting the dynamics of scientific activities: A diffusion-based network analytic methodology," in *2018 Annual Meeting of the Association for Information Science and Technology*, Vancouver, Canada, 2018.
- [38] T. L. Griffiths, M. I. Jordan, J. B. Tenenbaum, and D. M. Blei, "Hierarchical topic models and the nested chinese restaurant process," in *Advances in Neural Information Processing Systems*, 2004, pp. 17-24.
- [39] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, *Ethically Aligned Design: Prioritizing Human Wellbeing with Autonomous and Intelligent Systems*, First ed. IEEE, 2019.
- [40] M. Z. Wu, Yi, "Hierarchical topic tree: A hybrid model incorporating network analysis and density peaks searching," presented at the The 18th International Conference on Scientometrics & Informetrics, Leuven, Belgium, 2021.
- [41] M. Bastian, S. Heymann, and M. Jacomy, "Gephi: An open source software for exploring and manipulating networks," *Proceedings of International AAAI Conference on Web and Social Media*, vol. 8, pp. 361-362, 2009.
- [42] M. E. Newman, "Modularity and community structure in networks," *Proceedings of the National Academy of Sciences*, vol. 103, no. 23, pp. 8577-8582, 2006.
- [43] L. Waltman, N. J. van Eck, and E. C. Noyons, "A unified approach to mapping and clustering of bibliometric networks," *Journal of Informetrics*, vol. 4, no. 4, pp. 629-635, 2010.
- [44] M. Krzywinski *et al.*, "Circos: An information aesthetic for comparative genomics," *Genome Research*, vol. 19, no. 9, pp. 1639-1645, 2009.
- [45] P. Kalluri, "Don't ask if artificial intelligence is good or fair, ask how it shifts power," *Nature*, vol. 583, no. 7815, pp. 169-169, 2020.
- [46] A. Z. Huq, "A Right to a Human Decision," *Va. L. Rev.*, vol. 106, p. 611, 2020.
- [47] T. C. King, N. Aggarwal, M. Taddeo, and L. Floridi, "Artificial intelligence crime: An interdisciplinary analysis of foreseeable threats and solutions," *Science and Engineering Ethics*, vol. 26, no. 1, pp. 89-120, 2020.
- [48] W. N. Price, S. Gerke, and I. G. Cohen, "Potential liability for physicians using artificial intelligence," *Jama*, vol. 322, no. 18, pp. 1765-1766, 2019.
- [49] L. Drew, "The ethics of brain-computer interfaces," *Nature*, vol. 571, no. 7766, pp. S19-S19, 2019.

- [50] Y. Zhang, X. Cai, C. V. Fry, M. Wu, and C. Wagner, "Topic Evolution, Disruption and Resilience in Early COVID-19 Research," *SSRN*, 2020, doi: 10.2139/ssrn.3675020.
- [51] B. L. Webber, S. Raghu, and O. R. Edwards, "Opinion: Is CRISPR-based gene drive a biocontrol silver bullet or global conservation threat?," *Proceedings of the National Academy of Sciences*, vol. 112, no. 34, pp. 10565-10567, 2015.
- [52] B. Green, G. Cunningham, A. Ekblaw, P. Kominers, A. Linzer, and S. P. Crawford, "Open data privacy," *Berkman Klein Center Research Publication*, no. 2017-1, pp. 17-07, 2017.