A Bibliometric Review Of Academic Collaboration In The Governance, Risk And Compliance Of Artificial Intelligence

David Lau Keat Jin[1], Ganthan Narayana Samy[2], Fiza Abdul Rahim[3], Mahiswaran Selvananthan[4], Nurazean Maarop[5] & Noor Hafizah Hassan[6]

[1,2,3,5,6]Advanced Informatics Department, Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Kuala Lumpur, 54100, Malaysia
[4]Perdana Department, Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Kuala Lumpur, 54100, Malaysia
davidkeat@graduate.utm.my

Article history

Received:
11 Sep 2023

Received in revised form:
25 Oct 2023

Accepted:
25 Oct 2023

Published online:
18 Dec 2023

*Corresponding author
davidkeat@graduate.utm.my

Abstract

As Artificial Intelligence becomes increasingly pervasive in its application, research efforts have shifted from the discourse on ethical principles to the mechanism of implementation that would eventually bring benefits to the world's population while minimizing its risks. The adoption of best practices in governance, risk management and compliance would not only promote these objectives but also foster greater adoption of the technology by nations of the world. As Artificial Intelligence models are dependent upon the data ingested and used by algorithms, the input from researchers of different nationalities has the potential to reduce bias, enhance interoperability, and facilitate the generation of accurate decisions. Moreover, formulation of universal legislations and standards that are applicable to all nations would spur compliance and acceptance of AI solutions by its potential users. Thus, this study aims to shed light on the level of international collaboration among academicians by analyzing the publications in the past decade. The results show that the major contributors in this domain are the USA, UK, and China. Also, most of the contributions are from selected academic institutions in those countries only. Hence, greater collaborations can be forged with notable researchers from these institutions by researchers and practitioners from other countries to ensure that the development and use of AI can benefit all mankind in our increasingly connected societies.

Keywords: *Artificial Intelligence, governance, risk management, compliance, collaboration*

1. Introduction

Artificial Intelligence (AI) systems are defined as systems that display intelligent behavior by analyzing their environment and taking actions with some degree of autonomy toward achieving specific goals often on par or exceed human intelligence [1]. As AI is dependent upon the analysis of data feed into the system and applying algorithmic manipulations on those data to generate the required output, there are bound to be imperfections in its operation [2]. Undesirable events as reported by media include racial bias [3], gender bias [4], discrimination in credit-score rating [5], inaccuracy in grading teachers [6] and examination results [7], as well as a the more catastrophic consequences of causing road accidents [8]. When

* Corresponding author. davidkeat@graduate.utm.my

such incidents occur, the affected parties would require explanations regarding the way in which the AI system produce its conclusion or results [9]. Depending on the model used, there may be many stages of algorithms applied to the data fed into the system which may not be possible to explain, let alone interpret it by folks that are not trained in the field [10]. Apart from that, if personally identifiable information are ingested by the system, such as in finance and medical profession, the system may not comply with contemporary legislations regarding data privacy and protection [11, 12]. All these issues require deliberations by the stakeholders, particularly the organizations that are involved in its development, deployment, use and maintenance [13]. Furthermore, the number of incidents and controversies related to AI is increasing as reported by organization for AI, Algorithmic, and Automation Incidents and Controversies (AIAAIC) as shown in Figure 1.

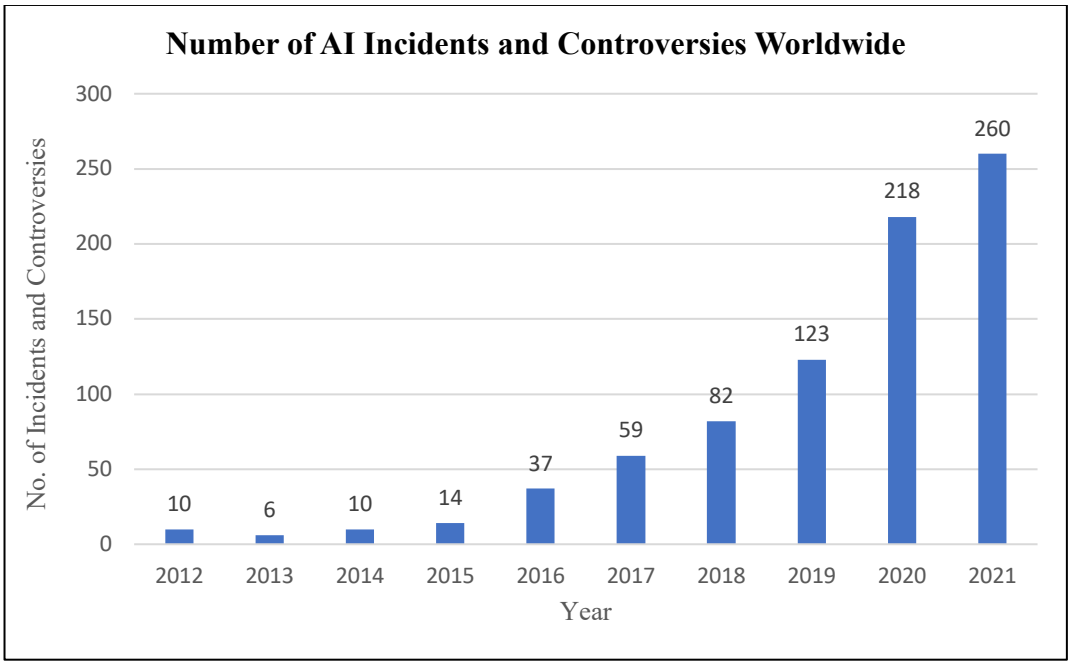

Figure 1. AI Incidendents and Controversies from 2012 till 2021 [14]

While the primary use of AI is to raise automation with the end results of elevating efficiencies and optimize the use of manpower, it can also be deliberately used for malicious purposes [15]. An example of such abuse is demonstrated by the use of Deepfakes to impersonate another person for scam [16]. When undesirable incidents occur due to unintentional or intentional use of AI, accountability issues arise as multiple stakeholders are involved in the approval, design, development, deployment, operation, maintenance and oversight of the system [17, 18]. This is exacerbated by the inscrutability of certain AI models which can be considered "black box" and the absence of any formal regulation or standard that could enforce compliance on responsible parties [19, 20].

2. Background

A review of extant articles related to ethical concerns due to AI was conducted by [21] and the author found that governance is the central theme that could affect all ethical issues raised by previous studies. In this regard, governance can be considered as: "a system of rules, practices, processes, and technological tools that are employed to ensure an organization's use of AI technologies aligns with the

organization's strategies, objectives, and values; fulfills legal requirements; and meets principles of ethical AI followed by the organization" [22]. The major goals of governance would be to ensure compliance and to manage risks [23]. This is the reason why the European Union (EU) AI Act specifies four levels of risks in considering the requirements for the use of AI by member countries [24, 25]. "AI risk management can drive responsible uses and practices by prompting organizations and their internal teams who design, develop, and deploy AI to think more critically about context and potential or unexpected negative and positive impacts. Understanding and managing the risks of AI systems will help to enhance trustworthiness, and in turn, cultivate public trust." [26]. Before AI became mainstream, GRC framework was previously proposed for Information System . Hence, in the integrated governance framework proposed by [27], risk management and compliance with regulations are also incorporated as illustrated in Figure 2.

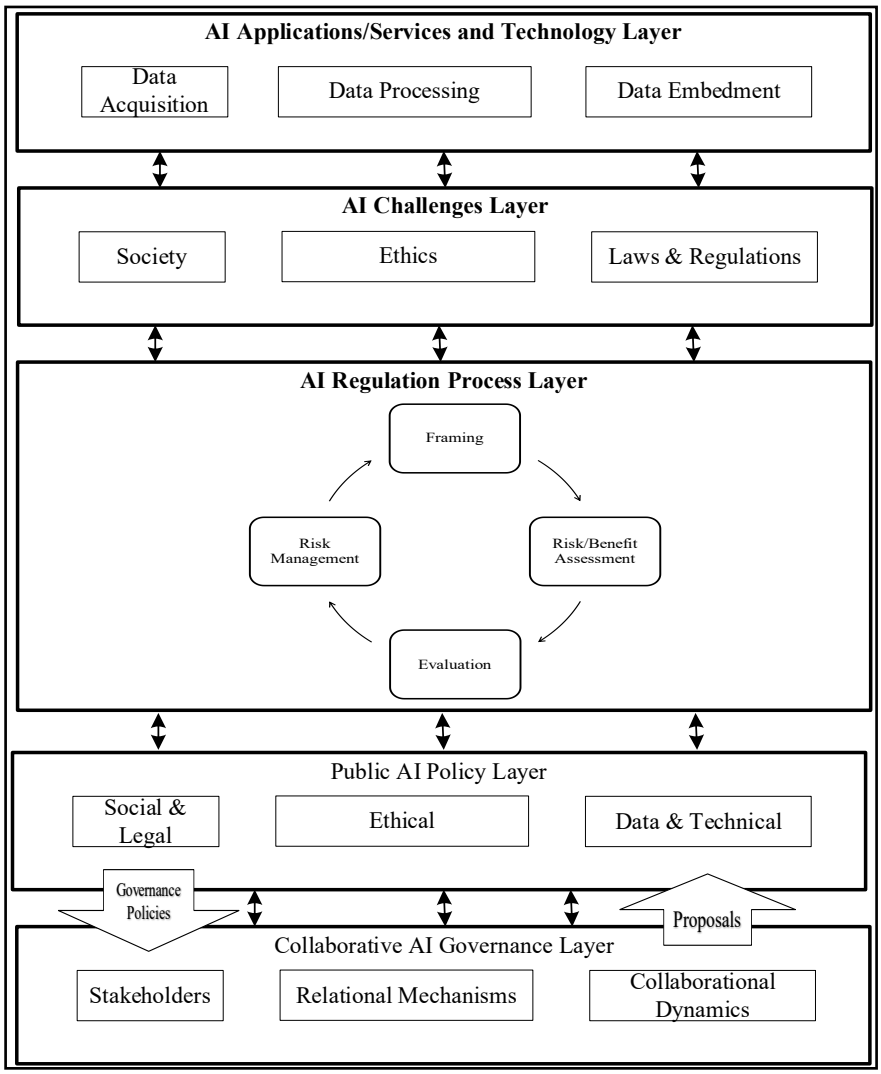

Figure 2. Integrated AI Governance Framework [27]

An international collaboration was manifested through "Montréal Declaration For A Responsible Development Of Artificial Intelligence" [28] in 2019 which agreed upon the ten principles as well-being, respect for autonomy, protection of privacy and intimacy, solidarity, democratic participation, equitable, diversity inclusion, caution, responsibility and sustainable development. In the following

year, Organization for Economic Cooperation and Development highlighted the principles for responsible stewardship of trustworthy AI as inclusive growth; sustainable development and well-being; human-centered values and fairness; transparency and explainability; robustness, security and safety; and accountability [29]. In reining over these concerns and considering the dependencies for AI related components in delivering a complete AI system, the World Economic Forum released their guidelines for AI procurement by governmental bodies [30]. As this progressive technology continues in its rapidly innovative path, there are bound to be novel issues that need international consultation and collaboration. Hence, it is timely for this study to provide an overview of the collaborative efforts undertaken by the academic communities. In line with this intention, Table 1 articulates the research questions and objectives of this study.

Table 1. Research Questions and Objectives

No.	Research Question	Research Objective
1.	What is the publication trend in the domain of governance, risk or compliance (GRC) in the past decade?	To validate the relevance of the research area.
2.	Who are the most cited authors in GRC approaches for AI?	To explore the relevance of the research domain and authors that contributed in the domain.
3.	Which countries' academic work are most frequently cited?	To ascertain the relevance of the countries' academic contribution.
4.	Which institutions produced the most academic articles?	To identify the most academically productive institutions so that future researchers can collaborate with the said institutions.
5.	What are the collaborative relationships among countries?	To illustrate the collaborative relationships among the countries.

3. Methodology

Bibliometrics is a statistical approach for assessing and analyzing the output of scientists, cooperation between universities, the effect of state-owned science funding on national research and development performance and educational efficiency [31]. In the past, this approach has been applied in the study of AI in various fields such as in operations environment [32], wastewater treatment [33], supply chain [34], tourism and hospitality [35] as well as renewable energy [36]. While there are various tools developed to process documents gathered for bibliometric analysis, [37] highlighted that *Bibliometrix* contains the more extensive set of techniques and suitable for practitioners through *Biblioshiny* [38]. With these rationales in place, the phases for this study are illustrated in Figure 3.

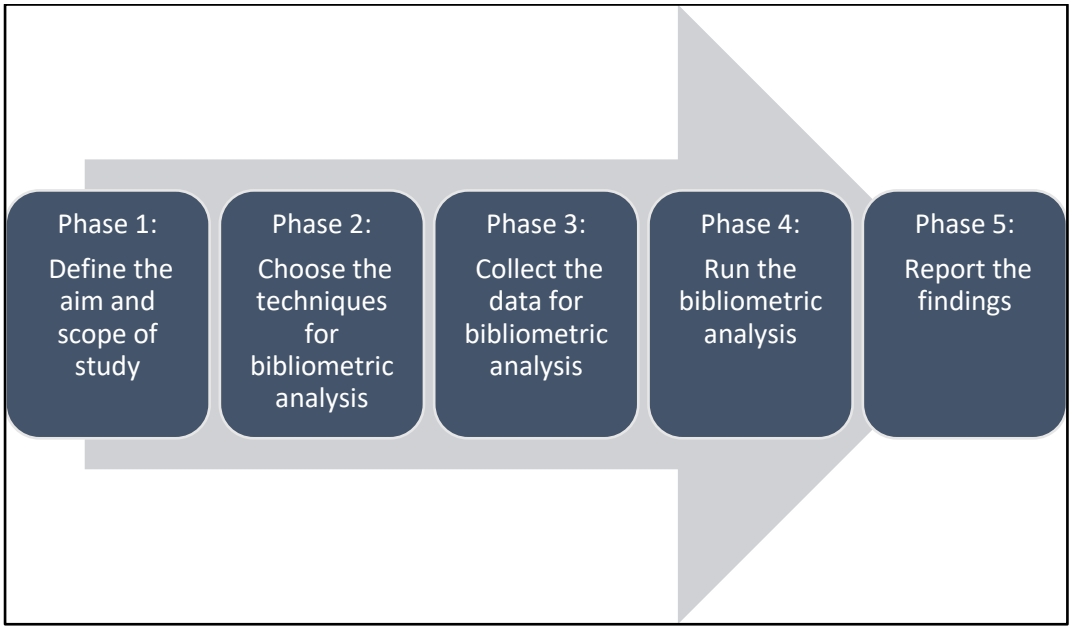

Figure 3. Phases of Bibliometric Analysis

In line with Phase 3 of Figure 3, Web of Science (WoS) and Scopus are the two databases used for data collection in this study which are accessible from Universiti Teknologi Malaysia's academic portal. WoS is an academic database that covers multiple disciplines and it is owned by Thomson & Reuters Corporation [39]. It is a source for high quality and reliable academic information and eventually became the main stream source for bibliometric analysis [40]. Comparatively, Scopus is the largest database of peer-reviewed journals [41] and it has been used by several other bibliometric studies [42]. Hence, these two databases are chosen for this study. Table 2 presents the method of searching from these two databases.

Table 2. Inclusion and Exclusion Criteria For Bibliometric Review

Attribute	**Inclusion Criteria**	**Exclusion Criteria**
Database	Scopus and WoS	Other databases
Search item	Title, abstract and author's keyword	Other metadata available
Search string	**Scopus** TITLE-ABS-KEY ((governance AND "artificial intelligence") OR ("risk management" AND "artificial intelligence") OR (compliance AND "artificial intelligence") not) AND PUBYEAR > 2012 AND PUBYEAR < 2024 AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk") OR LIMIT-TO (DOCTYPE , "cr")) AND (LIMIT-TO (LANGUAGE , "English")) **Web of Science**	Not applicable

Attribute	Inclusion Criteria	Exclusion Criteria
	The following string for topic is applied: (governance AND "artificial intelligence") OR ("risk management" AND "artificial intelligence") OR (compliance AND "artificial intelligence")	
Published period	2013 until 2023 for Scopus; 01/01/2013 until 31/08/2023 for Web of Science	Period not within the inclusion range
Reference type	Research Articles, Review Article, Books, Book Chapters, Conference Papers	Websites, Magazines, Non-academic reports
Language	English	Other languages

For this study, *EndNote* is used as the document management software while *RStudio*, *Biblioshiny* and *Excel* are used for data analysis. Prior to data analysis, the duplicate references obtained from WoS and Scopus databases are dropped. This is achieved by running the following commands in *RStudio* command line prior to importing the file named mergedb.xlsx in *Biblioshiny*.

```
S = convert2df("scopus.bib", dbsource = "scopus", format = "bibtex")
W = convert2df("wos.bib", dbsource = "isi", format = "bibtex")
Database = mergeDbSources(S, W, remove.duplicated = TRUE)
dim(Database)
install.packages("openxlsx")
write.xlsx(Database, file = "mergedb.xlsx")
```

4. Result

The total number of references obtained from WoS and Scopus are 2,225 and 1,338 respectively. The total number of duplicates removed is 472 resulting in the totally distinct reference as 3,091. The number of publication is on an increasing trend from the list of references extracted from both databases as evident from Figure 4. It should be noted that the number of publication for the year 2023 is until August 2023 and does not represent the full-year amount. The total number of references from 2013 until 2023 is given in Appendix A.

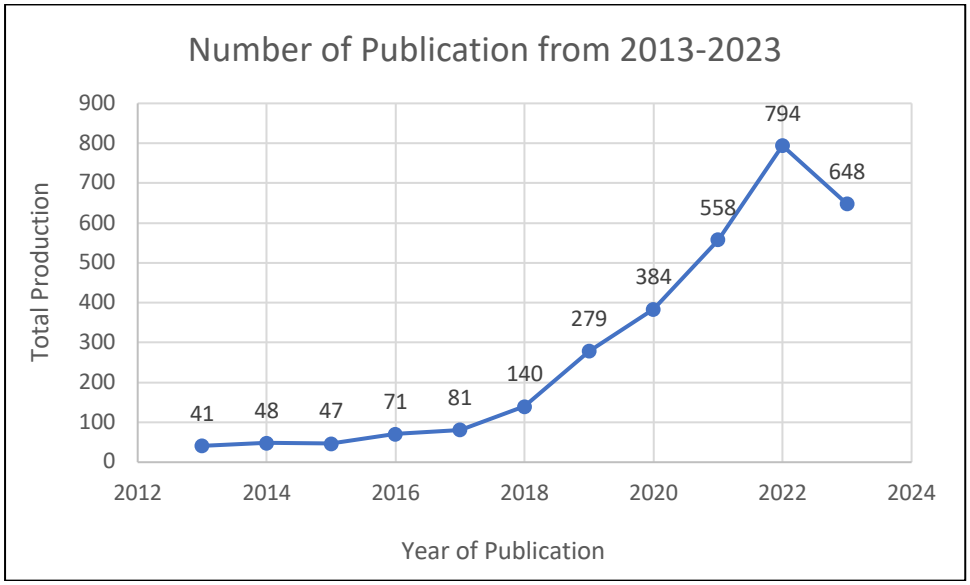

Figure 4. Number Of Publications For The Past 10 Years

The authors cited most frequently for their academic work are illustrated in Figure 5. Future researchers can determine if they should collaborate with these authors based on their experiences. The full list of authors is given in Appendix B.

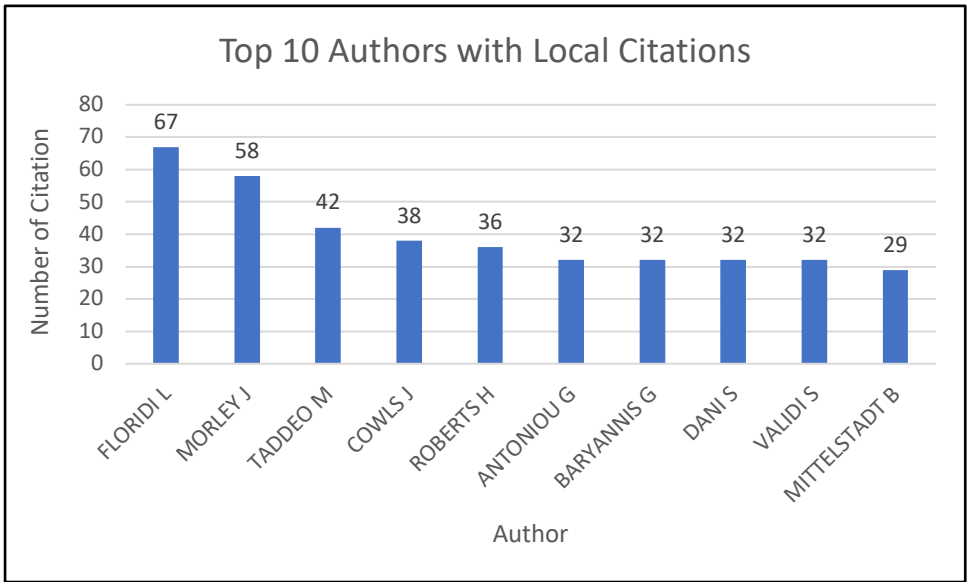

Figure 5. Top Authors by Number of References

Based on the authors' country membership, the countries with the most citations are given in Figure 6. It represents the importance of their authors' work as it is cited by other researchers. The full list of countries with number of citations is given as Appendix C.

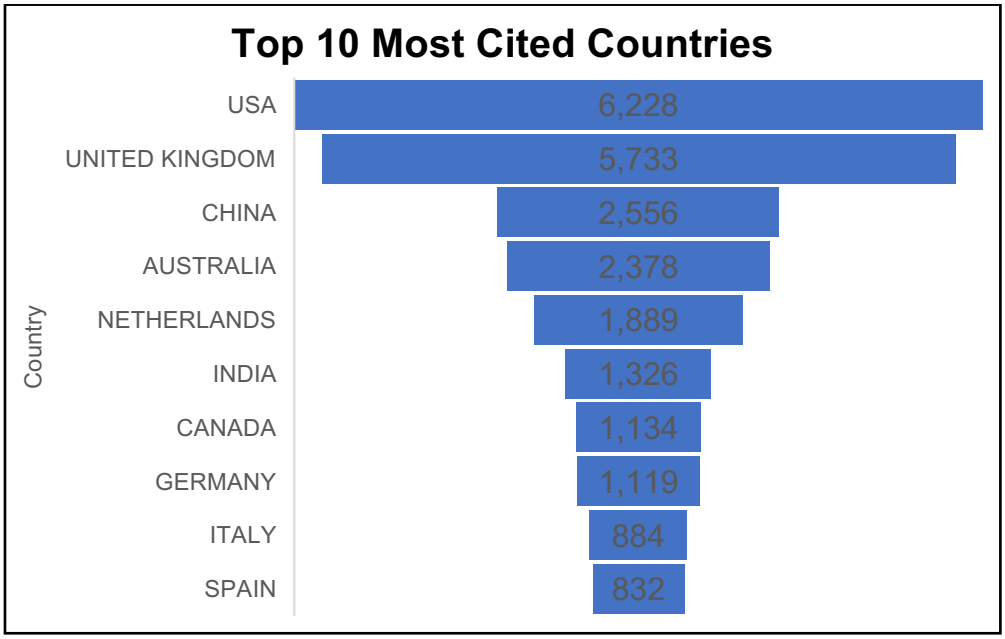

Figure 6. Top 10 Countries Based on Citations

The collaborative relationship between countries is illustrated in Figure 7. The thick lines indicate that the major research cooperations are between the countries of Canada and Europe, Australia and Europe, US and China, US and Europe. Hence, it can be concluded that the there are awareness and willingness by the academic community within each continent to collaborate to achieve greater effectiveness in GRC for AI. The full list of collaborative relationships is given in Appendix D.

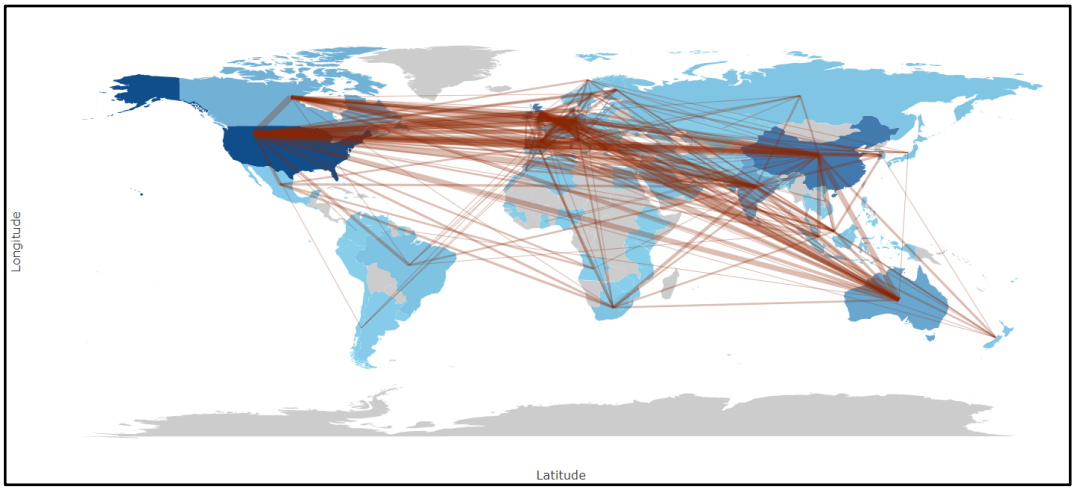

Figure 7. Collaboration Of Academic Researches In The World

The institutions that contributed the most in terms of academic articles are depicted in Figure 8. From the figure, it is clear that academic institutions are still at the forefront of research in this domain compared to business entities. University of Oxford produced the greatest number of studies in this domain followed by University of Toronto and University of Cambridge. There is a large gap between the number of publications between the University of Oxford and University of Toronto which is as much as 78 making the University of Oxford way ahead from other institutions in terms of publication. The complete list of publications by each affiliation is given in Appendix E.

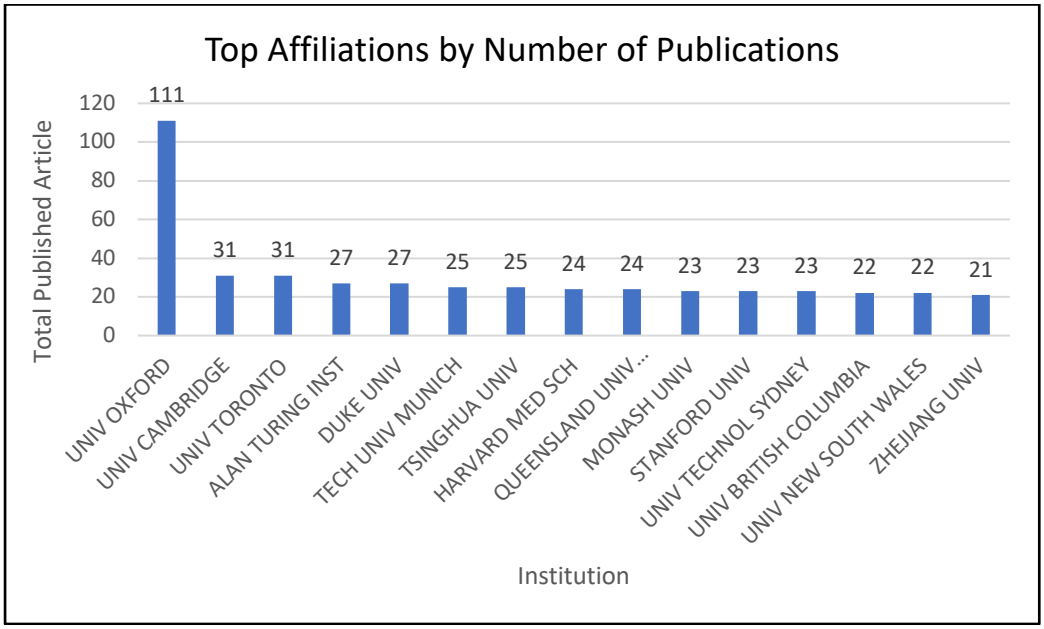

Figure 8. Top 10 Affiliations By Number of Publications

In addition to answering the research questions, the proximity of the terms governance, risk and compliance is gauged by the occurrence of the keywords either individually or concurrently in the field of title, author's keyword or abstract as presented in Table 3. From the result, the occurrence of the keyword is highest in the abstract of the references, while it occurred the least for simultaneous usages in all the said fields. While the concurrent use of all GRC terms in all the 3 fields are as much as 16% of the references examined, this is still equivalent to 515 of the references. Arguably, the result confirms that the three terms are closely linked in the discourse regarding governance, risk management or compliance with regards to AI technology. The complete list of references is given in Appendix F.

Table 3. Keyword Occurrence Analysis For GRC

Occurrence of the terms (governance, risk or compliance)	Number of Occurrence	Percentage of References
Title only	798	25.82%
Keyword only	1013	32.77%
Abstract only	2655	85.89%
Title and Keyword	2561	82.85%
Title and Abstract	2363	76.45%
Keyword and Abstract	848	27.43%
Title, Abstract and Keyword	515	16.66%

5. Discussion

Based on Figure 4, GRC approaches for AI is gaining attention by the research community. In addition, the authors that are most impactful in this field can be seen in Figure 5. Also, the countries that are at the forefront of this research domain are USA, United Kingdom, China, Australia and Netherlands with more than 100 publications for the past decade as illustrated in Figure 6. While future researches can look for the researchers from these countries due to their experiences, it also implies that more studies should be implemented by the countries that devoid in

publications particularly for field-dependent researches like actual application of AI in different environment such as in North Asia, South East Asia and South America. This is reinforced by the collaborative relationships demonstrated in Figure 7.

Interestingly, the most prolific universities in this research domain are situated in USA, Europe and Australia as evident from Figure 8. Institutions from other countries would do well to collaborate with these institutions to ensure that their countries are not left behind in terms of GRC development as AI becomes pervasive throughout the globe. Last but not the least, GRC is a progressive domain of research as new models of AI with more ubiquitous use cases are offered to the public. It would be interesting to see standards similar to IT governance like testing, audit requirements and certifications to be ratified through international collaboration among researchers.

6. Conclusion

Just as multi-disciplinary approaches and consultations with different stakeholders are required to obtain a holistic impact of AI systems, cross-country and cross-cultural cooperation are paramount to ensure that AI serves the needs of all humans and not to a privileged few or people from certain geographic regions per se. Such approaches stimulate the growth in development and adoption of the technology as well as enhance its adaptability for different geographical and cultural needs. This is in line with the vision of human-centered AI as espoused by various researchers [43-47]. Apart from these benefits, multi-ethnic cooperation underpins efforts to reduce bias from decision-support systems embedded with algorithms of AI. Furthermore, the accuracy of output generated can be assessed by researchers from different countries operating AI in diverse environment. This is true for autonomous systems which are used for self-driving vehicles and clinical systems used for diagnosis of diseases. Similar studies can be conducted to gauge the levels of cooperation in specific domain of AI to uncover its application, challenges, and research directions.

Acknowledgments

The author would like to thank the Malaysian Public Services Department that sponsored the research.

References

[1] Group, I.C.A., *IEEE guide for terms and concepts in intelligent process automation*. 2017.
[2] Guan, H., L. Dong, and A. Zhao. (2022). Ethical Risk Factors and Mechanisms in Artificial Intelligence Decision Making. *Behavioral Sciences, 12*(9). doi:10.3390/bs12090343
[3] Jeff Larson, S.M., Lauren Kirchner, Julia Angwin. *How We Analyzed the COMPAS Recidivism Algorithm*. 2016 Aug 25, 2023]; Available from: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.
[4] Dastin, J. *Amazon scraps secret AI recruiting tool that showed bias against women*. 2018 Aug 23, 2023]; Available from: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G.
[5] Plotnikova, V., M. Dumas, and F. Milani. (2020). Adaptations of data mining methodologies: A systematic literature review. *PeerJ Computer Science, 6*, e267.
[6] Guo, J., et al. (2021). An ai-application-oriented in-class teaching evaluation model by using statistical modeling and ensemble learning. *Sensors, 21*(1), 241.
[7] BBC. *A-levels and GCSEs: How did the exam algorithm work?* 2020 Aug 23, 2023]; Available from: https://www.bbc.com/news/explainers-53807730.
[8] NTSB. *Tesla Crash Investigation Yields 9 NTSB Safety Recommendations*. 2020 Aug 23, 2023].
[9] Lomas, D.A.S. (2022). The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: a comparative analysis of US, EU, and UK regulatory frameworks.
[10] Kaur, D., et al. (2022). Trustworthy artificial intelligence: a review. *ACM Computing Surveys (CSUR), 55*(2), 1-38.
[11] Arima, H. and S. Kano. (2021). Integrated Analytical Framework for the Development of Artificial Intelligence-Based Medical Devices. *Therapeutic Innovation & Regulatory Science, 55*(4), 853-865. doi:10.1007/s43441-021-00292-x

[12] Lore, F., et al. (2023). An AI framework to support decisions on GDPR compliance. *Journal of Intelligent Information Systems*, 28. doi:10.1007/s10844-023-00782-4
[13] Yurrita, M., et al. *Towards a multi-stakeholder value-based assessment framework for algorithmic systems*. in *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 2022.
[14] AIAAIC. *AI Index Report*. 2023 September 18, 2023]; Available from: https://spectrum.ieee.org/state-of-ai-2023.
[15] Blauth, T.F., O.J. Gstrein, and A. Zwitter. (2022). Artificial Intelligence Crime: An Overview of Malicious Use and Abuse of AI. *Ieee Access, 10*, 77110-77122. doi:10.1109/access.2022.3191790
[16] Verdoliva, L. (2020). Media forensics and deepfakes: an overview. *IEEE Journal of Selected Topics in Signal Processing, 14*(5), 910-932.
[17] Burton, S., et al. (2020). Mind the gaps: Assuring the safety of autonomous systems from an engineering, ethical, and legal perspective. *Artificial Intelligence, 279*, 103201.
[18] Santoni de Sio, F. and G. Mecacci. (2021). Four responsibility gaps with artificial intelligence: Why they matter and how to address them. *Philosophy & Technology, 34*, 1057-1084.
[19] Kouroupis, K. (2022). The AI Act in light of the EU Digital Agenda: A critical approach. *Journal of Data Protection and Privacy, 5*(3), 216-229.
[20] Erdélyi, O.J. and J. Goldsmith. *Regulating artificial intelligence: Proposal for a global solution*. in *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. 2018.
[21] Ashok, M., et al. (2022). Ethical framework for Artificial Intelligence and Digital technologies. *International Journal of Information Management, 62*, 102433. doi:https://doi.org/10.1016/j.ijinfomgt.2021.102433
[22] Mäntymäki, M., et al. (2022). Defining organizational AI governance. *AI and Ethics, 2*(4), 603-609. doi:10.1007/s43681-022-00143-x
[23] Cihon, P. (2019). Standards for AI governance: international standards to enable global coordination in AI research & development. *Future of Humanity Institute. University of Oxford*, 340-342.
[24] Mökander, J., et al. (2022). Conformity assessments and post-market monitoring: a guide to the role of auditing in the proposed European AI regulation. *Minds and Machines, 32*(2), 241-268.
[25] Kop, M. *Eu artificial intelligence act: The european approach to ai*. 2021. Stanford-Vienna Transatlantic Technology Law Forum, Transatlantic Antitrust ….
[26] Tabassi, E. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0).
[27] Wirtz, B.W., J.C. Weyerer, and B.J. Sturm. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration, 43*(9), 818-829.
[28] Declaration, M., *Montréal declaration for a responsible development of artificial intelligence*. 2019.
[29] Yeung, K. (2020). Recommendation of the council on artificial intelligence (OECD). *International legal materials, 59*(1), 27-34.
[30] WEF. *Guidelines for AI Procurement*. 2019 31 May 2023]; Available from: https://www3.weforum.org/docs/WEF_Guidelines_for_AI_Procurement.pdf.
[31] McBurney, M.K. and P.L. Novak. *What is bibliometrics and why should you care?* in *Proceedings. IEEE international professional communication conference*. 2002. IEEE.
[32] Dhamija, P. and S. Bag. (2020). Role of artificial intelligence in operations environment: a review and bibliometric analysis. *The TQM Journal, 32*(4), 869-896.
[33] Zhao, L., et al. (2020). Application of artificial intelligence to wastewater treatment: A bibliometric analysis and systematic review of technology, economy, management, and wastewater reuse. *Process Safety and Environmental Protection, 133*, 169-182.
[34] Riahi, Y., et al. (2021). Artificial intelligence applications in supply chain: A descriptive bibliometric analysis and future research directions. *Expert Systems with Applications, 173*, 114702.
[35] Knani, M., S. Echchakoui, and R. Ladhari. (2022). Artificial intelligence in tourism and hospitality: Bibliometric analysis and research agenda. *International Journal of Hospitality Management, 107*, 103317.
[36] Zhang, L., J. Ling, and M. Lin. (2022). Artificial intelligence in renewable energy: A comprehensive bibliometric analysis. *Energy Reports, 8*, 14072-14088.
[37] Moral-Muñoz, J.A., et al. (2020). Software tools for conducting bibliometric analysis in science: An up-to-date review. *Profesional de la Información, 29*(1).
[38] Aria, M. and C. Cuccurullo. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of informetrics, 11*(4), 959-975.
[39] Falagas, M.E., et al. (2008). Comparison of PubMed, Scopus, web of science, and Google scholar: strengths and weaknesses. *The FASEB journal, 22*(2), 338-342.
[40] Donthu, N., et al. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of business research, 133*, 285-296.
[41] Norris, M. and C. Oppenheim. (2007). Comparing alternatives to the Web of Science for coverage of the social sciences' literature. *Journal of informetrics, 1*(2), 161-169.
[42] Schiuma, G., et al. (2023). Research constituents and authorship patterns in the knowledge management research and practice: A bibliometric analysis. *Knowledge Management Research & Practice, 21*(1), 129-145.
[43] Ben, S. (2020). Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-centered AI Systems. *Acm Transactions on Interactive Intelligent Systems, 10*(4), 31. doi:10.1145/3419764
[44] Garibay, O.O., et al. (2023). Six Human-Centered Artificial Intelligence Grand Challenges. *International Journal of Human-Computer Interaction*, 47. doi:10.1080/10447318.2022.2153320
[45] Nagitta, P.O., et al. *Human-centered artificial intelligence for the public sector: The gate keeping role of the public procurement professional*. in *3rd International Conference on Industry 4.0 and Smart Manufacturing (ISM)*. 2021. Upper Austria Univ Appl Sci, Hagenberg Campus, Linz, AUSTRIA: Elsevier Science Bv.
[46] Rodriguez-Doncel, V., et al. *Introduction: A Hybrid Regulatory Framework and Technical Architecture for a Human-Centered and Explainable AI*. in *International Workshop on AI Approaches to the Complexity of Legal Systems (AICOL) / 3rd Workshop on Explainable and Responsible AI in Law (XAILA) at 33rd International Conference on Legal Knowledge and Information Systems (JURIX)*. 2020. Electr Network: Springer International Publishing Ag.
[47] Shneiderman, B. (2020). Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy human-centered AI systems. *ACM Transactions on Interactive Intelligent Systems, 10*(4). doi:10.1145/3419764