

Artificial Intelligence and Open-Source Models - The New Frontier of Systemic Risk

Mpho Mokgathi

Department of Commercial Accounting
College of Business and Economics
University of Johannesburg
Johannesburg, Republic of South Africa
mmokgathi@uj.ac.za

Omobolade Stephen Ogundele

Department of Commercial Accounting,
College of Business and Economics
University of Johannesburg
Johannesburg, Republic of South Africa
omoboladeo@uj.ac.za

Abstract

This study examined the growing intersection between artificial intelligence (AI) and systemic risk with attention to the open-source models as an emerging source of vulnerability in complex systems. The objective of the study is to systematically review and map global research trends on AI-driven systemic risk using a quantitative bibliometric approach. Data were sourced from the Scopus database, yielding a total of 218 relevant publications. The study employed VOS viewer to conduct keyword co-occurrence, overlay, network, and density visualisation, alongside descriptive analysis of publication trends, subject areas, funding sources, and country contributions. The findings revealed a significant upward trend in publications, especially from 2020 onwards, indicating increasing scholarly interest in AI-related systemic risk. Network analysis identifies four major research clusters: i) machine learning and financial systems ii) decision-making, ethical technology, and cybersecurity iii) risk assessment and deep learning iv) financial markets and electronic trading. Artificial intelligence emerges as the domain research hotspot, while areas such as ethical technology remain underexplored. The United States, China, and the United Kingdom are leading contributors, with computer science, social sciences, and finance as the most prominent subject areas. The study further highlights that while AI enhances innovation and efficiency, it simultaneously introduces systemic vulnerabilities through model opacity and security risks, especially within open-source ecosystems. The paper contributes to the literature by providing a comprehensive mapping of the intellectual structure of AI and systemic risk research. It also offers policy-relevant insights, emphasizing the need for robust governance frameworks.

Keywords

Artificial Intelligence, Bibliometric Analysis, Open-Source Models, Systemic Risk, Systematic Review

1. Introduction

The swift proliferation of open-source models and artificial intelligence (AI) has ushered in a transformative era, reshaping economies and industries while simultaneously introducing unprecedented systemic risks. Open-source AI platforms, such as Hugging Face, have democratised access to pre-trained models, which enables widespread collaboration and innovation. Also, the open-source AI models, while fostering innovation and collaboration, are

specifically exposed to some security threats such as membership inference, model inversion, backdoor attacks, and data leakage. These vulnerabilities arise from the inherent transparency of open-source systems, which can expose sensitive data or compromise system integrity (Al-Kharusi et al., 2024). However, accessibility and transparency come with major vulnerabilities, including security risks such as data leakage, backdoor attacks, and model inversion, which can compromise system integrity and sensitive data (Al-Kharusi et al., 2024; Kathikar et al., 2023). The open-source nature of these models amplifies the challenge, as their widespread adoption often outpaces the development of robust security measures (Kathikar et al., 2023). Moreover, the integration of artificial intelligence into significant sectors such as healthcare, finance, and infrastructure has heightened concerns about cascading systemic risks, where failures in one domain could propagate across interconnected systems (Zhou & Papageorgiou, 2025; Ozili, 2026).

The adoption of AI in the financial sector has redefined risk topologies, introducing novel challenges such as dynamic network rewriting, algorithmic homogeneity, and emergent multi-agent behaviour that exacerbate structural fragility (Zhou & Papageorgiou, 2025). The opacity of AI models and their convergence tendencies could lead to unforeseen vulnerabilities, such as market distortion, herding behaviour, and accelerated decision-making during a crisis, which amplify stress and destabilise financial systems (Peterson, 2024; Ameur et al., 2026). These risks are compounded by a lack of standardised governance frameworks, as the regulatory landscape for AI remains underdeveloped and fragmented (Mueck et al., 2025). As AI continues to evolve, the duality of its potential serving as both a source of systemic risk and a tool for innovation underscores the need for ethical oversight, robust governance frameworks, and adaptive regulatory mechanisms to harness its benefits while mitigating unintended consequences for sustainable development. This necessitates a paradigm shift in risk management. Also, this includes the development and adoption of probabilistic risk assessment models and the integration of advanced mitigation strategies such as differential privacy and adversarial training (Al-Kharusi et al., 2024; Maertens et al., 2024).

The systemic risks associated with open-source AI extend beyond technical vulnerabilities to include broader regulatory and societal challenges. The European Union's AI Act, for instance, categorises certain AI systems as high-risk, imposing stringent requirements for data quality, transparency, and human oversight, adoption of probabilistic risk assessment models, and the integration of advanced mitigation (Obradov & Pavković 2025). AI and open-source models represent a double-edged sword in systemic risk management. While they offer transformative potential, their adoption introduces vulnerabilities, especially in financial systems. The motivation for this study stems from the increased adoption of AI in financial systems, which has introduced new layers of interconnectedness and complexity, amplifying systemic risk. The need for a systematic review of the risks posed by open-source AI models is pivotal. Such a review would not only identify existing vulnerabilities but also propose frameworks for security, governance, and ethical oversight. By addressing these challenges, stakeholders can work towards building a resilient ecosystem that balances innovation with accountability and safety.

1.1 Objectives

The objective of the study is to systematically review and analyse the systemic risks associated with the integration of artificial intelligence and open-source models, with associated systemic risk.

2. Literature Review

Empirical studies are increasingly demonstrating that the integration of AI has implications for systemic risk. Peterson (2024), using analytical and case-based evidence, found that adoption of AI amplifies financial stability risks through heightened cybersecurity exposure, increased market independence, and heightened cybersecurity exposure, all of which can accelerate crisis transmission. Similarly, Ameur et al. (2026) provide empirical insights from financial institutions showing that while AI enhances efficiency in credit scoring, fraud detection, and portfolio optimisation, the opacity of machine learning models and the tendency toward model convergence introduce systemic vulnerabilities that are difficult to monitor. Exploring simulation techniques, O'Halloran and Nowaczyk (2019) empirically evaluated regulatory interventions and demonstrated that although mechanisms such as collateralisation reduce resolution costs, they do not necessarily diminish systemic risk and may even be redistributed unevenly across financial participants. Further empirical modelling by Zhou and Papageorgiou (2025) revealed that AI-driven systems, especially in decentralised finance, contribute to systemic fragility through synchronised algorithmic behaviour, network complexity, and endogenous contagion effects. Collectively, these findings confirm that AI not only transforms financial operations but also reconfigures the transmission and structure of systemic risk.

Empirical evidence also highlights that open-source AI models significantly expand the systemic risk landscape by introducing governance, security, and scalability challenges. Kathikar et al. (2023), through a large-scale vulnerability assessment of over 110,000 models and 29,000 associated repositories, found that a substantial proportion of open-source AI systems contain high-severity vulnerabilities, posing risks to the broader digital and financial ecosystem. Complementing this, Eiras et al. (2024) employed a staged analytical framework to empirically assess generative AI development and conclude that while open-source models drive innovation and accessibility, they also increase exposure to misuse, model exploitation and systemic spillover if not properly governed. From a regulatory perspective, Obradov and Pavković (2025) evaluated compliance mechanisms under the EU AI Act and proposed a verification mechanism that integrates technical tools in assessing dataset quality, robustness, and transparency, thereby increasing the reliability of AI systems. Mueck et al. (2025) further provide empirical evidence and institutional analysis indicating that regulatory classifications of high-risk and general-purpose AI models are vital in mitigating systemic exposure, although open-source exemptions present ongoing challenges. Overall, the empirical literature suggests that the convergence of AI and open-source ecosystems creates a complex, interconnected risk environment, reinforcing the need for evidence-based regulatory and technological interventions in managing systemic risk effectively.

3. Methods

This study used the quantitative method. Data was downloaded from the Scopus database. The quantitative approach is defined as numerical data collected through observation that describes and interprets the phenomena represented by those observations (Chand, 2025). Since its introduction in 2004, the Scopus database has quickly become a strong alternative database and is now considered the largest source for quality peer-reviewed publications (Alonso-Álvarez, 2024). Statistical data were analysed using content analysis, wherein conference papers, journals, articles, and books are included in this data. This method was seen as best for reaching the study goal, which is to map world study trends on Artificial Intelligence and Systemic Risks. The aim is to spot the mind setup, theme growth, and new holes in knowledge in this field. A cross-sectional approach was used, since the data were collected once in a single setting.

4. Data Collection

As mentioned in the methodology section, the study used quantitative bibliometric analysis using Scopus as a database and analysed it using VOSviewer version 1.6.19 and Scopus database analysis. The search strings were “Artificial Intelligence” AND “Systemic Risk”. The total number of documents published is 218. Regarding analysis, the following was done:

VOSviewer:

Co-occurrence of All Keywords,

- Counting Method – Full counting
- Minimum occurrences are standard at 7
 - o 1820 are total keywords, and 36 met the threshold
 - o The following keywords were removed from the 34 selected keywords: Article, which occurred 11 times and has a total link strength of 57. Humans, which occurred 11 times and has a total link strength of 68. Banking, which occurred 8 times and has a total link strength of 39. COVID-19, which occurred 7 times and has a total link strength of 37. AI, which occurred 7 times and has a total link strength of 11, and lastly, diabetic retinopathy, which occurred 7 times and has a total link strength of 28. This is because the study focused on Artificial Intelligence and systemic risk research. This reduced the number of keywords that met the threshold to 29, which were used in the analysis.

Scopus database analysis:

The 218 documents were used to do the analysis. The following was done:

1. Publication per year
2. Funding Sponsor
3. Publication by Subject Area
4. Publication by Country

5. Results and Discussion

5.1 Numerical Results

VOSviewer highlighted the four clusters (1, 2, 3, & 4). They are categorised as follows: The one painted red is Cluster 1. Cluster 1 includes the following keywords: machine learning, forecasting, financial risks, finance, and learning systems. Cluster 2 is green, and it consists of the following keywords: decision making, ethical technology, risks management, systemic risks, information systems, and cybersecurity. Cluster 3 is painted blue with the following keywords: artificial intelligence, risk assessment, risk analysis, deep learning, and risk factor. Lastly, Cluster 4 is painted yellow and has the following keywords: financial market, commerce, investment, and electronic trading.

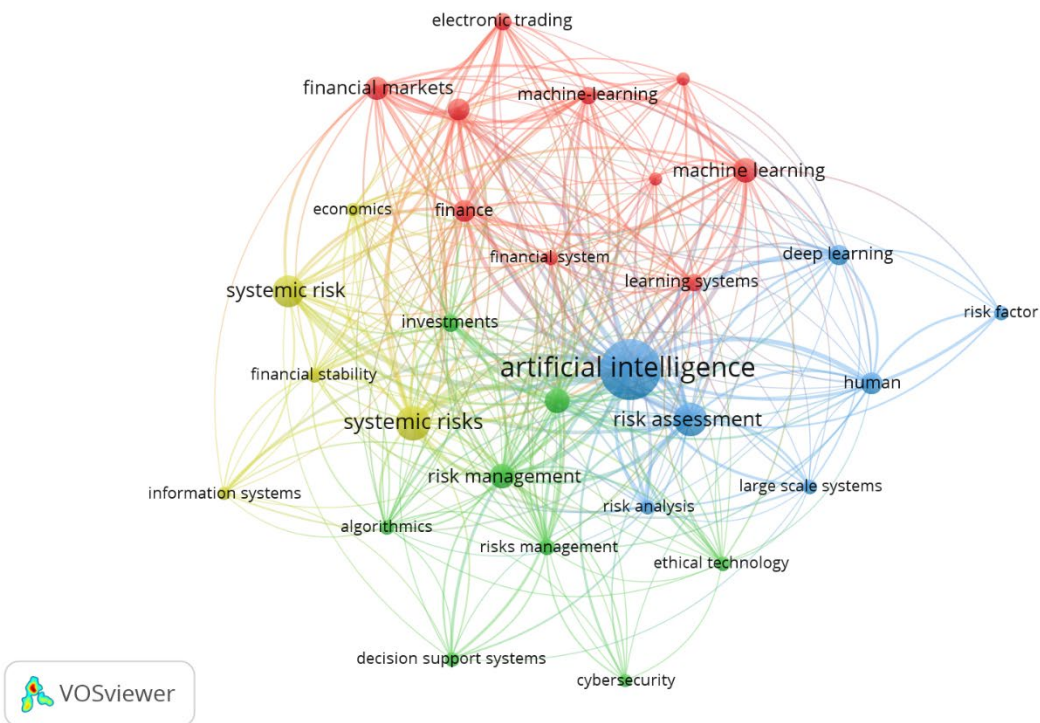


Figure 1. Network Visualisation of All Keywords

Source: Authors' Compilations

Figure 1 above shows a network visualisation of all the research keywords. This figure shows that artificial intelligence is the most frequently used keyword since it has a bigger node in the future. Artificial intelligence is in cluster 3; it has 29 links, a total link strength of 378, and has occurred 137 times. This is followed by risk assessment, in cluster 3, with 29 links, a total link strength of 190, and has occurred 41 times. This is followed by systematic risks in cluster 4, with 27 links, a total link strength of 195, and has occurred 46 times. This is followed by machine learning in cluster 1, with 26 links, a total link strength of 120, and has occurred 23 times. Lastly, decision-making in cluster 2, with 28 links and a total link strength of 127, has occurred 23 times.

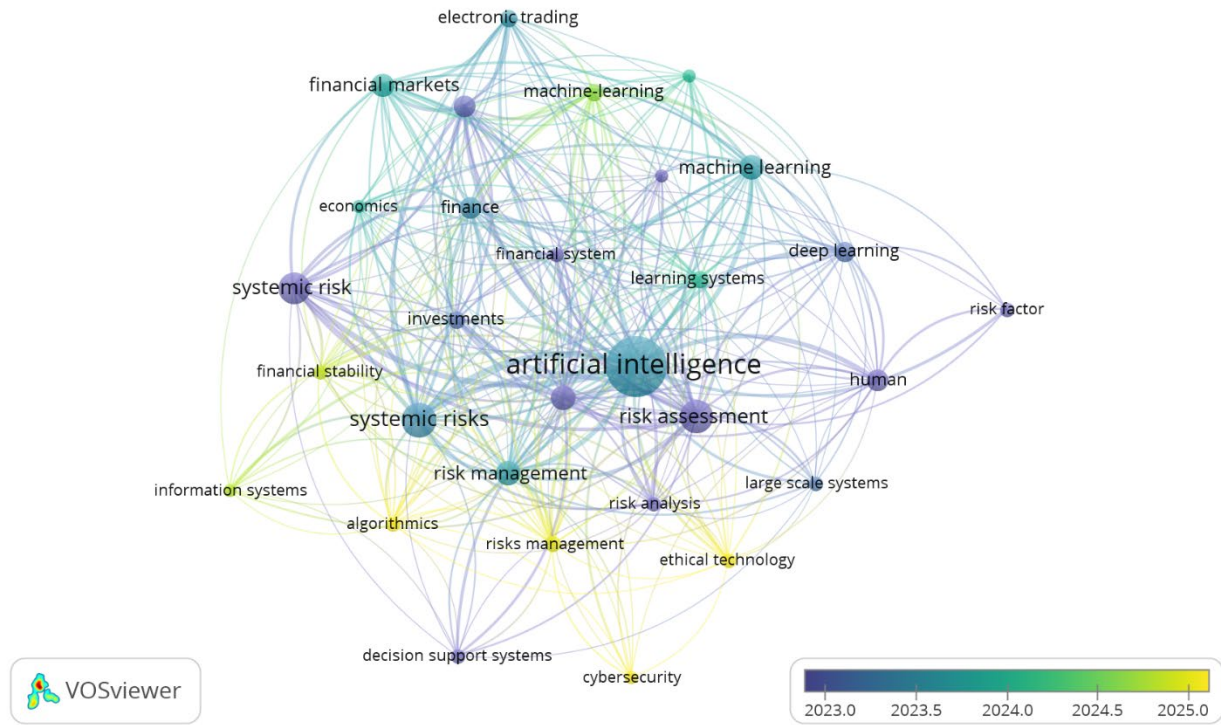


Figure 2. Overlay Visualisation

Source: Authors' Compilations

Figure 2 above depicts an overlay visualisation of all keywords. Artificial intelligence, on cluster 3, has 29 links and 378 total link strength, an occurrence of 137, and has an average publication year of 2023.54. This is an established research trend and was mostly used in 2023. This is followed by risk assessment, in cluster 3, with 29 links, a total link strength of 190, which has occurred 41 times and has an average publication year of 2022.71. This is followed by systematic risks in cluster 4, with 27 links, a total link strength of 195, which has occurred 46 times and has an average publication year of 2023.46. This is followed by machine learning in cluster 1, with 26 links, a total link strength of 120, which has occurred 23 times, and has an average publication year of 2023.65. Lastly, decision-making in cluster 2, with 28 links and a total link strength of 127, has occurred 23 times and has an average publication year of 2022.61.

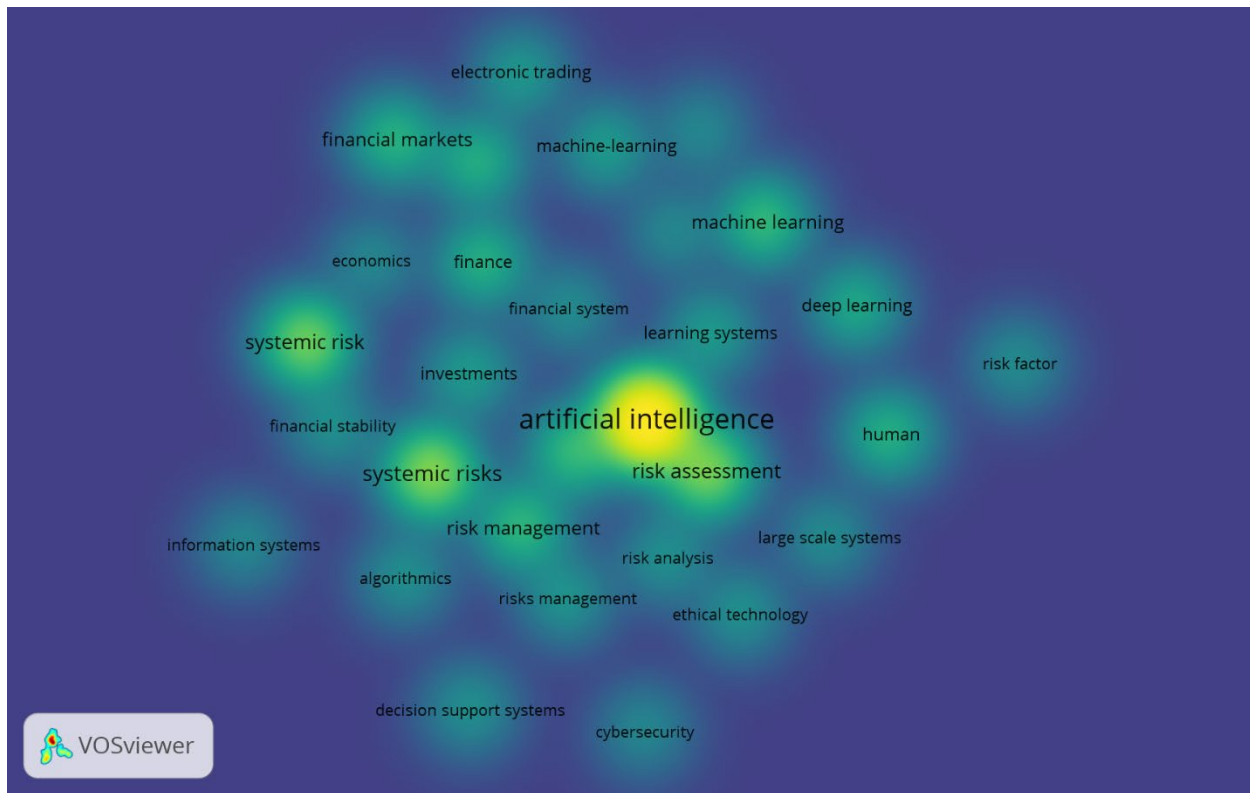


Figure 3. Density Visualisation
Source: Authors' Compilations

Figure 3 above shows that the hotspot regarding the topic at hand is artificial Intelligence. It shows that many people are writing about it. This is followed by Systemic risks, risk assessment, decision-making, machine learning, risk management, financial markets, information systems, and risk analysis, which are emerging research trends. Ethical technology is under-researched.

The following tables were drafted using the Scopus dashboard. Below is Figure 4, indicating the number of publications per year, followed by a discussion section.

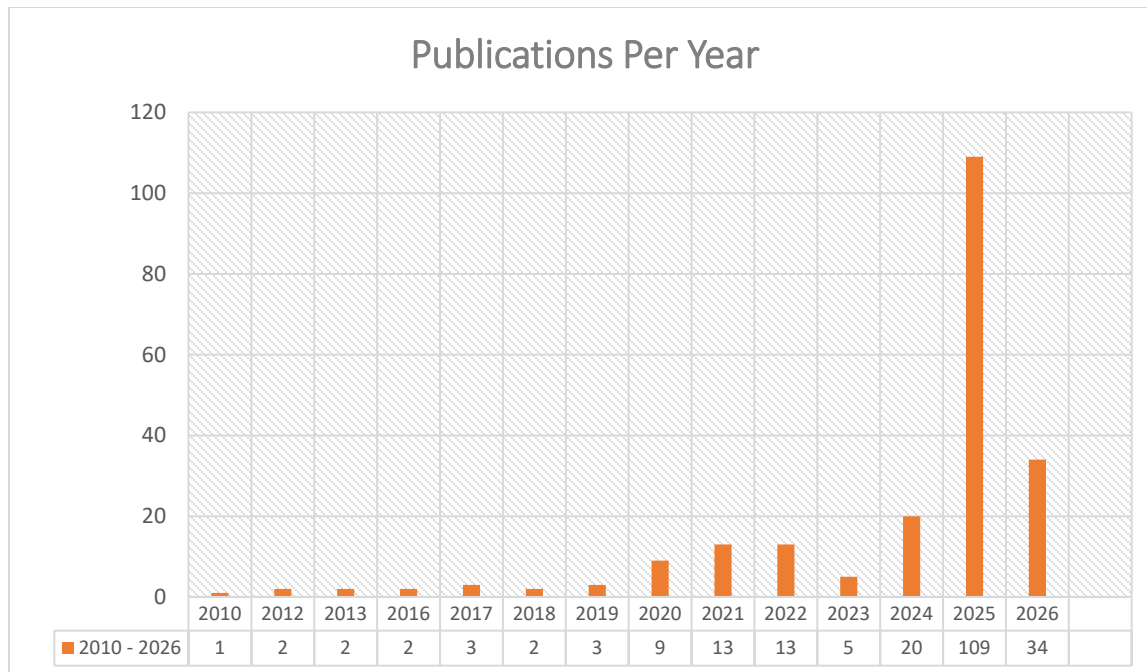


Figure 4: Number of Publications Per Year

Figure 4 shows the annual distribution of relevant publications from 2010 to 2026. It shows an uptrend of publications, which started peaking from 2020 with 9 publications, and the highest was in 2025 with 109 publications. 34 published documents were recorded on Scopus in 2026. However, the publications can increase since the year has not yet ended. Generally, the number of publications started well and increased over the years.

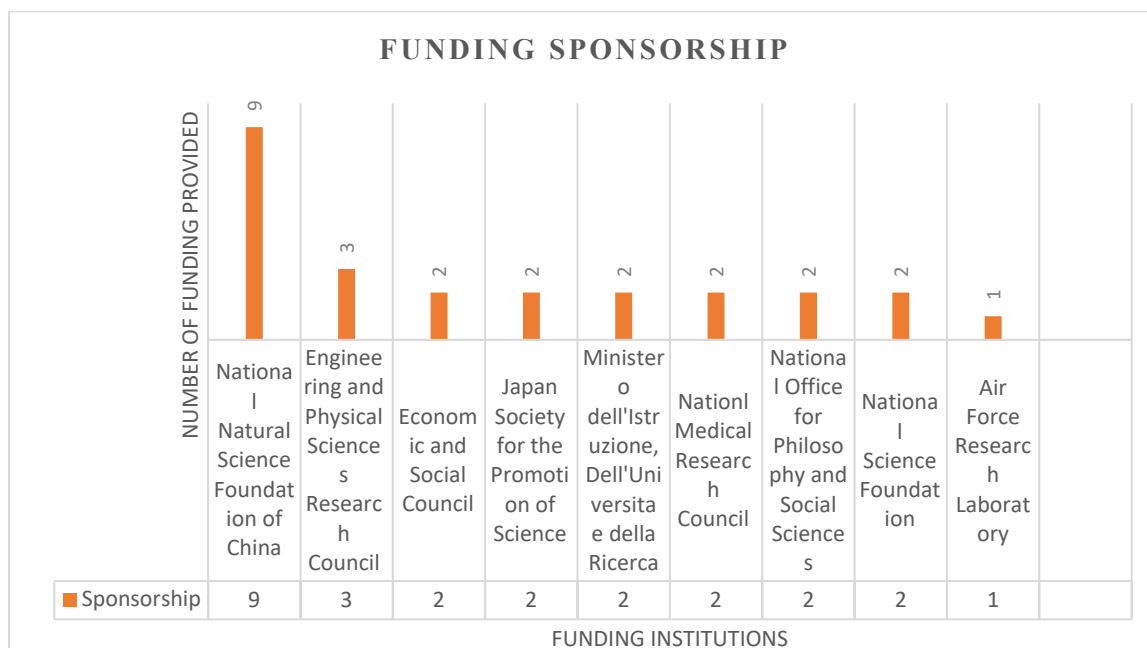


Figure 5: Funding Sponsorship

Figure 5 above indicates six institutions that funded Artificial Intelligence research across the world. The institution that funded the most research based on this study is the National Natural Science Foundation of China, with 9

publications. Then followed by the Engineering and Physical Science Research Council, with 3 publications. Then followed by the institutions like the Economic and Social Council, the Japan Society for the Promotion of Science, Ministero dell'Istruzione, Dell'Universita e della Ricerca, the National Medical Research Council, the National Office for Philosophy and social science, and the National Science Foundation, with 2 publications each. Lastly, the Air Force Research Laboratory has 1 publication.

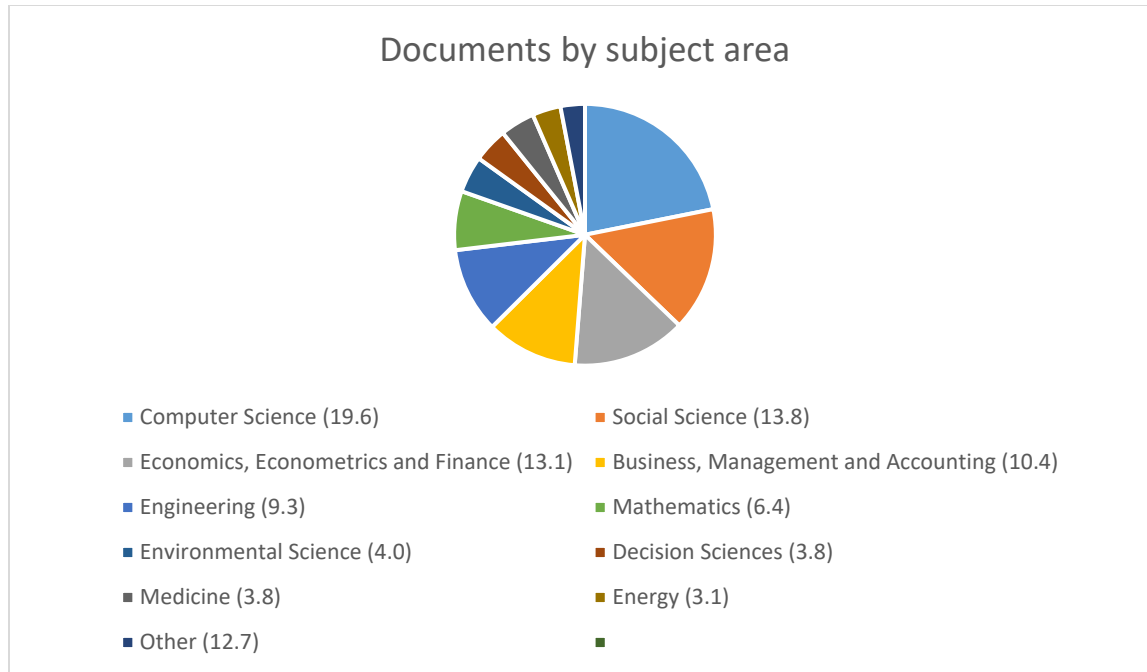


Figure 6. Pie chart: Publications by Subject Area

The pie chart above (Figure 6) depicts publications by subject type. Ten major subjects contribute to the body of knowledge as far as artificial intelligence is concerned. The biggest contributor is Computer Science with 19.6% of publications on scopus database. Then followed by the following subject areas: Social Science with 13.8% of publications, Economics, Econometrics, and Finance, with 13.1% of publications, Business, Management and Accounting with 10.4% of publications, Engineering with 9.3% of publications, Mathematics with 6.3% of publications, Environmental Science with 4.0% of publications, Decision sciences with 3.8% of publications, Medicine with 3.8% of publications, and lastly Energy with 3.1% of publications.

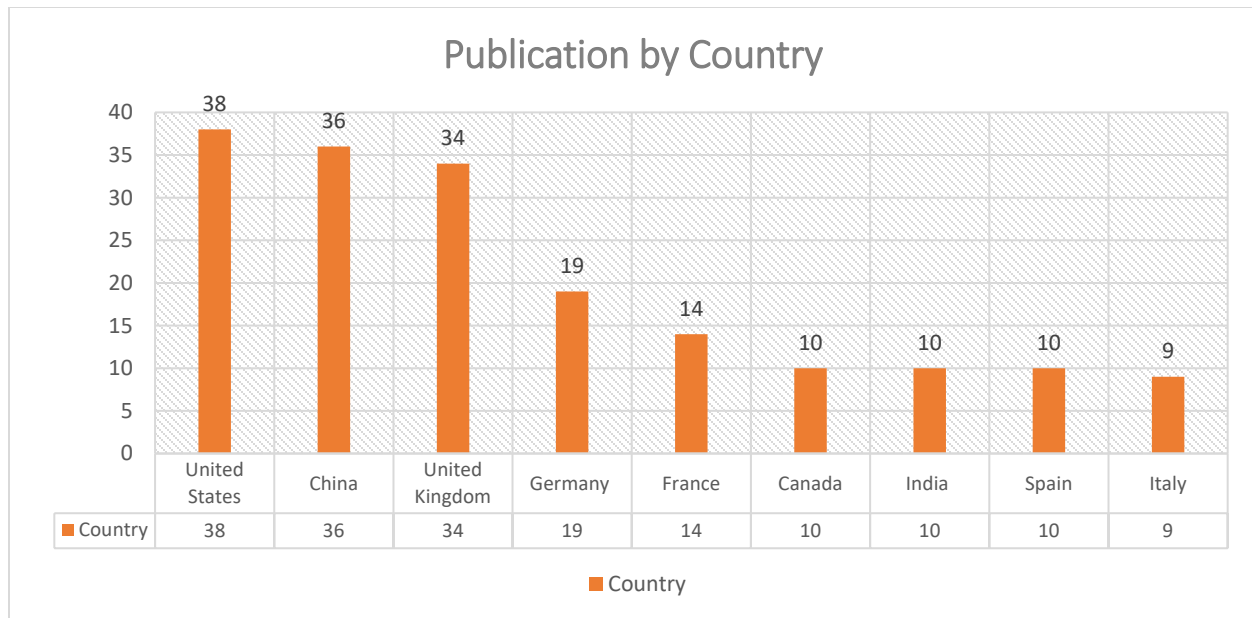


Figure 7: Publications by Country

Figure 7 above indicates the publications by country. The major contributor is the United States with 38 publications, followed by China with 36 publications, then the United Kingdom with 34 publications, then Germany with 19 publications, followed by France with 14 publications. Canada, India, and Spain all share 10 publications each. Lastly, Italy has 9 publications.

5.3 Proposed Improvements

While the bibliometric analysis provides a useful overview of the Artificial Intelligence and systemic risk research landscape, several methodological and presentational improvements are recommended. The search strategy using only "Artificial Intelligence" AND "Systemic Risk" may be too restrictive; incorporating related terms such as "machine learning," "deep learning," "neural networks," and "cascading failures" would likely capture a broader and more representative set of publications. Future work should expand the analytical toolkit beyond co-occurrence analysis to include bibliographic coupling, co-citation analysis, and factorial mapping, which could uncover deeper intellectual structures and research fronts. Lastly, integrating a qualitative validation component, such as expert interviews, case studies, or in-depth thematic analysis of highly cited papers, would complement the quantitative findings and provide richer contextual insights into emerging research gaps and ethical considerations.

6. Conclusion

The study successfully mapped global research trends at the intersection of artificial intelligence and systemic risk using quantitative bibliometric analysis of 218 Scopus documents. The findings confirm a sharp upward trajectory in publications since 2010, with artificial intelligence emerging as the dominant hotspot, while ethical technology remains significantly under-researched. The United States, China, and the United Kingdom lead in scholarly output, with computer science as the primary contributing discipline. The key unique contribution lies in the systematic identification of four distinct research clusters: machine learning and finance, decision-making and cybersecurity, risk assessment and deep learning, and financial markets and electronic trading, providing researchers with a clear structural map of the field. All stated objectives were fulfilled, and the study highlights an urgent need for future work on the ethical dimensions of AI-driven systemic risk management.

References

- Al-Kharusi, Y., Khan, A., Rizwan, M., and Bait-Suwailam, M. M. "Open-source artificial intelligence privacy and security: A review," *Computers*, Vol. 13, No. 12, p. 311, 2024. <https://doi.org/10.3390/computers13120311>
- Alonso-Álvarez, P. "Exploring research quality and journal representation: A comparative study of African Journals Online, Scopus, and Web of Science," *Research Evaluation*, Vol. 33, p. rvae057, 2024.

- Ameur, H. B., Casarin, R., Caporin, M., Ftiti, Z., and Maillet, B. B. "Towards a better understanding of financial and economic systems' complexities: Some new evidence coming from artificial intelligence, machine learning and big data advanced technologies," *Annals of Operations Research*, Vol. 357, pp. 1–10, 2026. <https://doi.org/10.1007/s10479-025-07025-5>
- Chand, S. P. "Methods of data collection in qualitative research: Interviews, focus groups, observations, and document analysis," *Advances in Educational Research and Evaluation*, Vol. 6, No. 1, pp. 303–317, 2025.
- Eiras, F., Petrov, A., Vidgen, B., Schroeder, C., Pizzati, F., Elkins, K., Mukhopadhyay, S., Bibi, A., Purewal, A., Botos, C., and Steibel, F. "Risks and opportunities of open-source generative AI," *arXiv preprint arXiv:2405.08597*, 2024.
- Kathikar, A., Nair, A., Lazarine, B., Sachdeva, A., and Samtani, S. "Assessing the vulnerabilities of the open-source artificial intelligence (AI) landscape: A large-scale analysis of the Hugging Face platform," *Proceedings of the 2023 IEEE International Conference on Intelligence and Security Informatics (ISI)*, pp. 1–6, 2023. <https://doi.org/10.1109/ISI58743.2023.10297271>
- Maertens, A., Antignac, E., Benfenati, E., Bloch, D., Fritsche, E., Hoffmann, S., Jaworska, J., Loizou, G., McNally, K., and Hartung, T. "The probable future of toxicology: Probabilistic risk assessment," 2024. <https://doi.org/10.14573/altex.2310301>
- Mueck, M., Gaie, C., and Gkikas, D. C. "Introduction to the European artificial intelligence act," in *European Digital Regulations*, pp. 53–90, Springer Nature Switzerland, Cham, 2025.
- Obradov, A., and Pavković, B. "Ensuring compliance of high-risk AI systems with the EU AI Act using open source libraries," *Proceedings of the 2025 IEEE Zooming Innovation in Consumer Technologies Conference (ZINC)*, pp. 54–59, 2025. <https://doi.org/10.1109/ZINC65316.2025.11103517>
- O'Halloran, S., and Nowaczyk, N. "An artificial intelligence approach to regulating systemic risk," *Frontiers in Artificial Intelligence*, Vol. 2, p. 7, 2019. <https://doi.org/10.3389/frai.2019.00007>
- Ozili, P. K. "Artificial intelligence (AI), financial stability and financial crisis," *International Encyclopedia of Business Management*, Vol. 3, pp. 591–595, 2026. <https://dx.doi.org/10.2139/ssrn.4970195>
- Zhou, Y., and Papageorgiou, O. "AI-DeFi systemic risk: A complex adaptive system framework," *Proceedings of the 2025 7th Conference on Blockchain Research & Applications for Innovative Networks and Services (BRAINS)*, pp. 1–9, 2025. <https://doi.org/10.1109/BRAINS67003.2025.11302949>

Biographies

Mpho Mokgathi holds Postgraduate Diplomas in Risk Management and Business Management from Mancosa. He enrolled for an MCom in Computer Auditing at the University of Johannesburg. He has been in the higher education sector for eight years; he started his teaching career at the University of Johannesburg in 2017, where he served as a Cost and Estimating and Marketing Management tutor for several years. Later, he joined Rosebank College as a lecturer and currently works at the Department of Commercial Accounting in the School of Management at the College of Business and Economics, University of Johannesburg, as an Assistant Lecturer. Mpho's research areas are IT risk management, the protection of personal data, cyber security, and computer auditing. He recently published research papers at the international conferences globally, held in Cape Town, South Africa, Bali, Indonesia, and Bangkok, Thailand. He has been actively engaged in community service since he was a student, and as a result, the University recognised his efforts. Mpho served as a project manager at the Enactus office, where he did a lot of work assisting the community, tutoring and training them in entrepreneurship. He taught many SME owners how best to leverage ICT in their businesses. Currently, he is involved in a Day Care project at Soweto. Mpho Mokgathi is a young, promising, and emerging academic who is well-rounded.

Dr Omobolade Stephen Ogundele is a Post-Doctoral Research Fellow in the Department of Commercial Accounting, College of Business and Economics, University of Johannesburg, South Africa. He holds Ph.D, M.Sc., and B.Sc in Accounting from Obafemi Awolowo University, Nigeria, and is a member of the Institute of Chartered Accountants of Nigeria. His teaching and supervision span undergraduate and postgraduate levels, with additional roles as an external examiner for postgraduate programmes. His research interests focus on corporate governance, corporate reporting, auditing, and the application of artificial intelligence in accounting and finance, especially within emerging economies. Dr Ogundele has published in peer-reviewed journals, conferences, and book chapters. He is actively involved in academic service as a reviewer for local and international journals and conferences. His work contributes to policy and practice by advancing accountability, transparency, and the use of innovative technology to enhance sustainable development.