Information Research - Vol. 31 No. 1 (2026)

Mapping the research landscape of artificial intelligence for knowledge discovery in innovation research

Yanyi Wu

DOI: https://doi.org/10.47989/ir31152435

Abstract

Introduction. Artificial intelligence (AI) is increasingly vital for knowledge discovery within innovation research (IR) and the science of science (SoS), yet its specific research landscape lacks systematic mapping. This study addresses this gap by providing a comprehensive bibliometric overview of AI's application for knowledge discovery in innovation research, aiming to structure the field and identify key trends.

Method. A bibliometric analysis was performed on 1,094 articles and reviews published between 2010 and 2024, retrieved from the Web of Science Core Collection. Data processing and visualisation employed VOSviewer and Bibliometrix.

Analysis. Descriptive statistics quantified publication growth and collaboration patterns. Network analyses mapped thematic structures, using keyword co-occurrence; identified intellectual foundations, through co-citation networks; and visualised current research frontiers through bibliometric coupling.

Results. Findings indicate exponential publication growth and high international collaboration, dominated by China and the USA. Key thematic clusters focus on AI methodologies (machine learning (ML), deep learning (DL), natural language processing (NLP)), innovation contexts (patent analysis, technology trends), and integrated science of science methods (bibliometrics, scientometrics). Intellectual foundations derive strongly from computer vision, sequence/topic modelling, and bibliometric tools.

Conclusion. This mapping structures the field, highlighting AI's profound integration as both a transformative tool for innovation analysis and an object of study within the science of science framework itself. It underscores the field's dynamism and provides a basis for future research on AI's impact and responsible application.

Introduction

Artificial intelligence (AI) has rapidly matured, creating a paradigm shift far beyond computational science. It fundamentally reshapes how we approach data analysis and knowledge generation across virtually all domains of human inquiry (Jarrahi et al., 2022). AI encompasses methods such as machine learning, deep learning, and natural language processing. These techniques possess unparalleled capabilities to process, interpret, and extract complex patterns from massive, high-dimensional datasets that overwhelm traditional analysis (Ayinaddis, 2025). Such prowess makes AI an exceptionally potent engine for knowledge discovery: the non-trivial process identifying valid, novel, potentially useful, and ultimately understandable data patterns (Dessimoz & Thomas, 2024). Applying AI to knowledge discovery accelerates insights, automates laborious analysis, and unveils previously hidden relationships. It promises to revolutionise scientific discovery and technological innovation speed and scope across diverse fields, from medicine and materials science to social systems and economic behaviour (Marino et al., 2023; Wang et al., 2023).

Innovation research constitutes a vibrant, essential component within the broader academic landscape, covering science, technology, and innovation studies, closely aligned with science of science (SoS) goals. Focus lies on understanding the multifaceted processes generating, developing, and integrating novelty into economic and societal structures (Fortunato et al., 2018). Innovation studies investigate complex interplay among various actors, including individuals, firms, research institutions, universities, and governments, exploring how their interactions shape technological trajectories, market dynamics, and societal progress (Lê & Schmid, 2020). As a data-intensive domain, innovation research frequently uses quantitative methods. Researchers analyse research and development patterns, evaluate innovation system performance, track technological diffusion, and map scientific and technological field evolution (Mariani et al., 2023a). Insights gleaned from innovation research are indispensable for informing policymaking, guiding strategic organisational decision-making, and fostering environments conducive to sustainable innovation (Howoldt, 2024).

The intersection of AI, knowledge discovery techniques, and innovation research is emerging as a critically important area, driven by the inherent data-rich nature of innovation processes. Innovation activities generate vast quantities of diverse data (spanning academic publications, patent filings, market data, collaborative networks, and organisational information), that present significant analytical challenges (Bogers et al., 2018). AI offers sophisticated solutions for extracting meaningful insights from these complex data streams. For example, machine learning algorithms can analyse patent databases to identify emerging technological trends (Chellappa et al., 2021); natural language processing can uncover hidden relationships and sentiment in research reports (Khurana et al., 2022); and network analysis, often enhanced by graph AI, can map and analyse complex innovation ecosystems (Maruccia et al., 2020). These AI-powered knowledge discovery applications enable innovation researchers to move beyond descriptive analysis to predictive modelling and prescriptive guidance, offering powerful new tools to understand, manage, and stimulate innovation.

Despite evident potential and growing instances where AI is applied for knowledge discovery within specific innovation research facets, a comprehensive, macro-level understanding regarding this interdisciplinary domain's overall research landscape remains notably underdeveloped. Existing literature offers valuable insights in applying particular AI techniques to discrete innovation research problems, such as using deep learning for technology forecasting, or AI for analysing R&D collaborations (Gama & Magistretti, 2025; Mariani et al., 2023b). However, no systematic study has yet mapped the collective structure, thematic evolution, key actors, or intellectual foundations for this converging field as a whole. Consequently, researchers entering or operating within this space lack a synthesised overview covering its historical development, current hotspots, dominant methodologies, influential contributors, and prevailing collaboration patterns. Identifying key research gaps and promising future directions effectively is therefore challenging.

To address this significant gap and contribute a much-needed systemic perspective, this study undertakes a comprehensive bibliometric mapping of the research landscape at the intersection of Artificial Intelligence for knowledge discovery in innovation research. Employing established bibliometric methodologies and visualisation tools, the research aims to provide a data-driven overview of this burgeoning field. Specifically, this study is guided by the following research questions:

> RQ1: What are the publication trends and growth patterns in this study domain? Which are the primary publication sources and leading contributors based on publication output?
>
> RQ2: What is the thematic structure of the field, including key research topics and AI techniques, and how have these themes evolved?
>
> RQ3: What are the key collaboration networks among authors and countries shaping knowledge production in this study area?

By executing this systematic bibliometric analysis, the study offers several important contributions to the academic understanding and practical application of AI in innovation research. Theoretically, it provides the first comprehensive mapping of this specific interdisciplinary landscape, clarifying its boundaries, revealing its underlying structure, and tracing its evolution, thereby enriching the theoretical understanding of how AI methods are being integrated into the study of innovation. This work also contributes directly to the science of science by illustrating AI's role in enabling advanced, data-intensive analysis of innovation, a fundamental component of scientific and technological systems. Practically, the detailed maps of key themes, influential actors, and collaboration patterns will serve as an invaluable resource for researchers seeking to identify potential research niches, forge collaborations, and benchmark their contributions, while also providing policymakers and practitioners with data-backed insights to guide strategic planning and investment in AI-enabled innovation analysis capabilities.

Methodology

Data source and search strategy

For this bibliometric mapping study, the Web of Science (WoS) Core Collection database was selected as the primary source of scholarly literature. This database was chosen because of its extensive coverage of high-impact journals across a wide range of scientific disciplines, its robust indexing capabilities, including author-supplied keywords and editorially curated terms, and its comprehensive collection of citation data, all of which are essential for conducting thorough bibliometric and network analyses. To systematically identify publications at the intersection of AI and specific knowledge discovery tasks within innovation research, a focused search strategy was developed. The strategy combined terms representing AI and its core techniques (such as machine learning and deep learning) with keywords denoting specific, data-driven knowledge discovery processes frequently applied in biological and medical research, and finally constrained these results to the broader innovation research domains.

The search was executed within the topic field, utilising the following specific query: TS = (("Artificial Intelligence" OR AI OR "Machine Learning" OR ML OR "Deep Learning" OR "Neural Network*" OR "Natural Language Processing" OR NLP) AND ("Knowledge Discovery" OR "Data Mining" OR "Pattern Recognition" OR "Topic Modeling" OR "Knowledge Extraction" OR "Insight Generation") AND ("Innovation Research" OR "Innovation Studies" OR "Technolog* Innovation" OR "Innovation Management" OR "Technology Management" OR "Technology Forecasting" OR

"Technology Prediction" OR "Tech Mining" OR "Patent Analysis" OR "Scientific Literature Analysis" OR Bibliometrics OR Scientometrics OR "Emerging Technolog*" OR "Technology Trend*")). This precise formulation aimed to retrieve documents where these key concepts were prominently discussed in titles, abstracts, or keywords. The initial execution of this query on the specified date yielded a total of 2,633 records.

Data filtering and refinement

A series of filtering and refinement steps were applied to curate a relevant and manageable dataset for analysis. First, to focus specifically on core research outputs, the document type was limited to *article* and *review*, which reduced the dataset to 1,588 publications. Second, given the rapid advancements in both AI and innovation research, and to capture the most relevant and contemporary trends, the publication period was restricted to articles published between January 2010 and December 2024, resulting in a dataset of 1,096 documents. Third, the language was restricted to English, the predominant language of scientific communication in these fields. These final steps yielded a refined dataset comprising 1,094 documents, which were subsequently downloaded in a suitable format (e.g., plain text file) for import into the bibliometric analysis software tools. The entire data filtering process is illustrated in Figure 1.

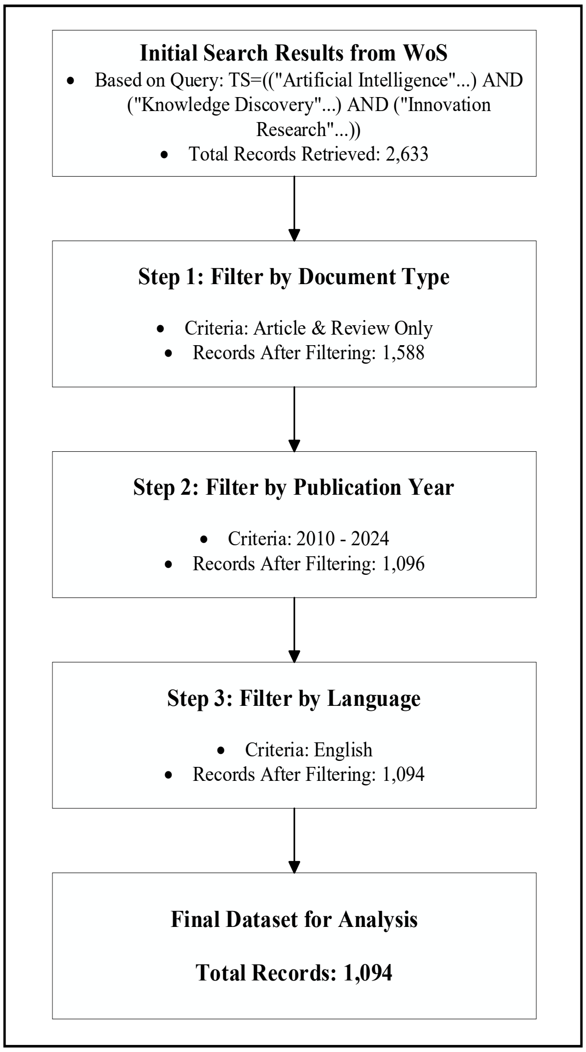

Figure 1. Flowchart of the literature filtering process

Bibliometric tools

Two primary software tools were employed to conduct the bibliometric analysis and visualise the research landscape: VOSviewer and Bibliometrix, including its associated Biblioshiny Web interface. VOSviewer is a sophisticated tool, specifically designed for creating and visualising bibliometric networks, enabling the exploration of relationships between various entities such as keywords, authors, institutions, and countries (van Eck & Waltman, 2010). It facilitates the construction of co-occurrence, co-authorship, and co-citation maps, providing spatial representations of the structure and dynamics of a research field. Bibliometrix, an R-package with a user-friendly Web interface (Biblioshiny), offers a comprehensive suite of quantitative methods for conducting descriptive statistical analyses and performing more advanced mapping techniques, such as thematic evolution analysis and bibliometric coupling, complementing the network visualisation capabilities of VOSviewer (Aria & Cuccurullo, 2017). These integrated tools provided the necessary functionalities to process the downloaded dataset and generate the statistical summaries, trends, tables, and network maps presented in this study.

Data analysis methods

Following the data collection and filtering process, the refined dataset was subjected to a comprehensive bibliometric analysis utilising the selected tools. The analysis was broadly divided into two main components: descriptive analysis and network analysis. The descriptive analysis involved calculating key metrics to quantify publication trends over time, identify the most prolific authors, influential institutions, leading countries, and frequently publishing journals within the field. It also included summarising information about keywords and highly cited documents to highlight core concepts and impactful research outputs. The network analysis component focused on revealing the structural relationships among different elements of the research landscape. It involved constructing and visualising various networks, including:

- co-occurrence networks of keywords to identify prominent research themes and their conceptual structure (Callon et al., 1983);
- co-authorship networks among authors, institutions, and countries to map collaboration patterns and social structures (Katz & Martin, 1997);
- co-citation networks of documents, authors, or journals to uncover the intellectual foundations and knowledge base of the field (Small, 1973);
- and bibliometric coupling networks at the document level to reveal similarities in research focus and identify current research frontiers (Kessler, 1963).

Findings

Descriptive analysis

Based on the carefully filtered dataset sourced from the WoS Core Collection, this descriptive analysis employs metrics generated by the Bibliometrix software package. Its purpose is to quantify publishing trends, pinpoint influential entities, and furnish fundamental insights into the structure and momentum of this burgeoning research landscape, thereby establishing the empirical foundation for the subsequent network analyses.

Brief information

An initial overview of the dataset, as summarised in Figure 2, reveals a collection of 1,094 documents published between 2010 and 2024. These publications originate from a substantial pool of 4,203 distinct authors affiliated with 451 different sources, encompassing journals and conference proceedings. Collectively, these works cite a vast body of literature, totalling 53,615 references. The average citation count per document stands at a notable 21.46, suggesting a domain where research findings are gaining scholarly recognition and impact. With an average document age of merely 3.43 years, the dataset prominently features contemporary research, indicative of a field experiencing vibrant and recent activity.

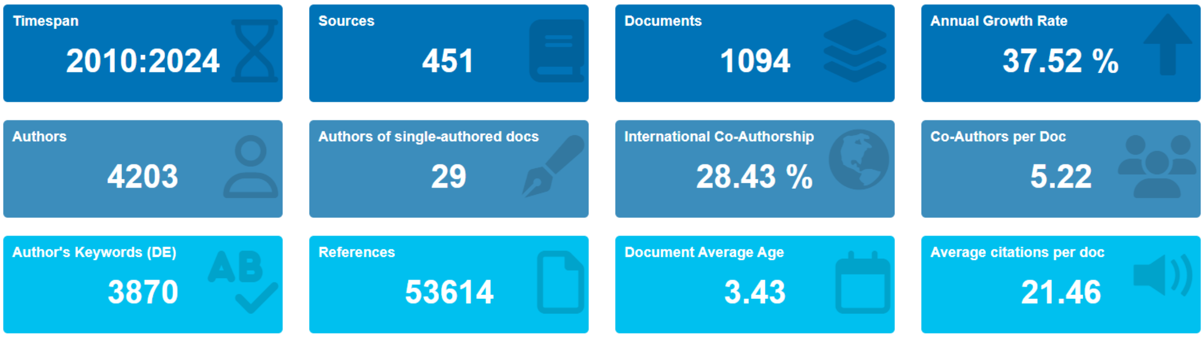

Figure 2. Overview of bibliometric data

A particularly striking statistic is the field's high annual growth rate, calculated at an impressive 37.52%, which unequivocally signals an exponential increase in scholarly output at the confluence of AI, knowledge discovery, and innovation research, especially evident in the latter part of the analysed period. Furthermore, the data paints a picture of a highly collaborative environment, evidenced by an average of 5.22 co-authors per document and an international co-authorship rate reaching 28.43%. Conversely, documents authored by a single individual constitute a marginal proportion (only twenty-nine documents), collectively underscoring the inherently interdisciplinary and globally interconnected nature of research within this specific area.

Annual scientific production

Figure 3 visually reinforces the extraordinary growth trajectory highlighted by the descriptive statistics. The annual scientific production curve exhibits a relatively gradual increase from 2010 through to approximately 2018. However, mirroring the field's rapid expansion, the number of publications per year displays a dramatically steeper incline from 2019 onwards, culminating in a significant peak in 2024. This pronounced acceleration in scholarly output during the more recent years strongly attests to the increasing strategic importance and practical adoption of AI-driven approaches for knowledge discovery tasks within innovation research contexts. It reflects not only advancements in AI methodologies, but also the growing availability of relevant data and the escalating interest from the research community in leveraging these capabilities to gain deeper insights into innovation processes and outcomes.

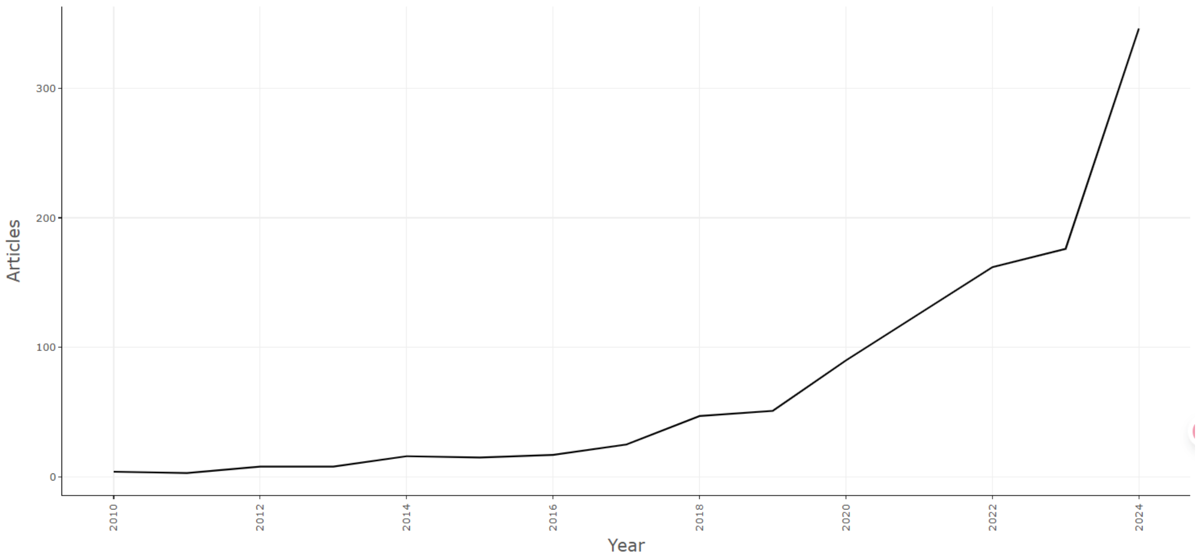

Figure 3. Annual scientific publications

Offering a more nuanced perspective on the field's scholarly impact, figure 4 tracks the average citations per document, per publication year, within the dataset. While the overall average citation rate suggests a field with considerable influence, the year-by-year averages reveal noticeable fluctuations, with peaks observed in earlier periods, specifically in 2013 and again in 2018. The subsequent decline in average citations for publications from the most recent years (post-2020) is a commonly observed pattern in fast-growing research domains, primarily because newer articles have had less time to accrue citations compared to their older counterparts (Waltman, 2016). Nevertheless, the presence of earlier citation peaks could signify periods during which particularly foundational or highly impactful studies were published, papers that continue to heavily influence the field and warrant closer examination in the analysis of highly cited documents later in this study.

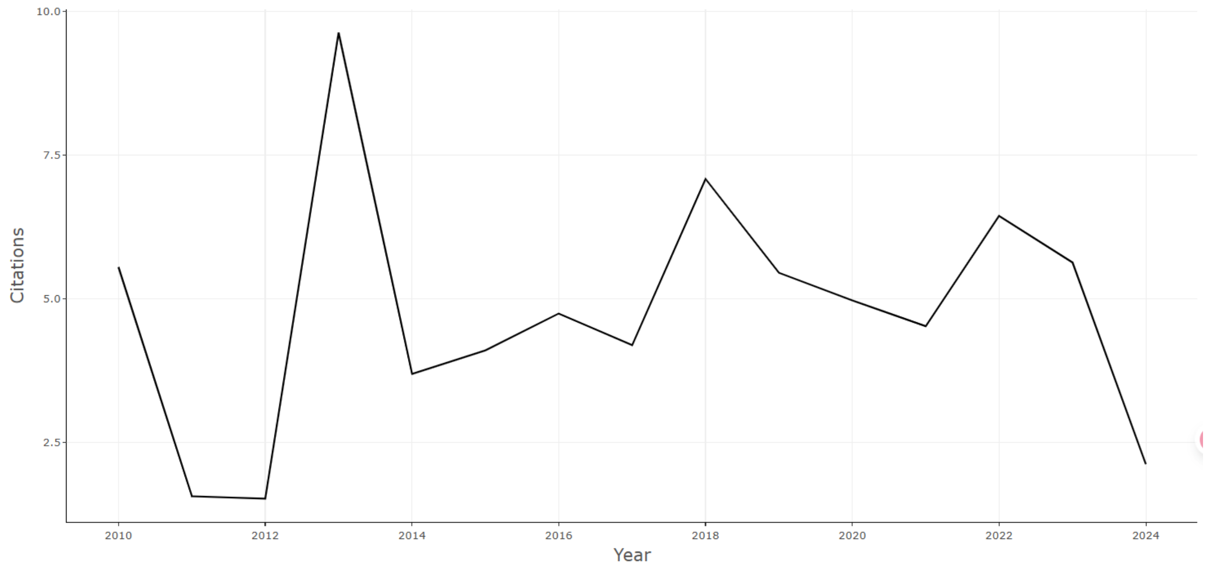

Figure 4. Average citations per year

Keywords

Valuable insights into the core concepts and dynamically evolving themes that define this study domain are gleaned from an analysis of author-supplied keywords. A review of the top keywords by frequency, presented in table 1, clearly identifies terms central to the field's discourse. High-frequency keywords such as *classification* (n=72), *model* (n=44), *system* (n=43), *network* (n=39), and methodological terms like *neural-networks* (n=30) and *algorithm* (n=27) prominently feature the AI techniques being applied. Concurrently, the frequent appearance of keywords denoting the application context, such as *management* (n=20) and the implicit presence of *innovation*, *technology*, and *forecasting* from the search strategy, underscores the embedding of these AI methods within innovation research. This consistent co-occurrence of computational techniques with innovation-related terminology compellingly affirms the interdisciplinary character of the field, where AI's analytical power is being directed towards understanding and analysing various facets of innovation processes and complex systems.

Words	Occurrences	Words	Occurrences
classification	72	framework	23
model	44	neural-network	23
system	43	design	20
network	39	management	20
prediction	35	optimization	20
performance	33	systems	20
science	33	recognition	19
neural-networks	30	selection	19
identification	28	networks	18
algorithm	27	artificial-intelligence	17

Table 1. Top keywords by frequency

The temporal trend analysis in figure 5 offers a dynamic view of the field's evolving focus. In the earlier years (roughly 2014-2018), foundational and more specific analytical techniques, such as *productivity* and *principal component analysis*, show initial activity. Entering the more recent period (approximately 2021 onwards), a noticeable concentration emerges relating to terms central to AI methodologies and their application frameworks. However, the most current hotspots, characterised by large circles positioned around 2022-2024, are dominated by core application-oriented terms like *system*, *network*, and *science*. This cluster strongly suggests that contemporary research is heavily invested in applying predictive modelling, analysing complex systems, developing robust models, and situating these efforts within a broader scientific context.

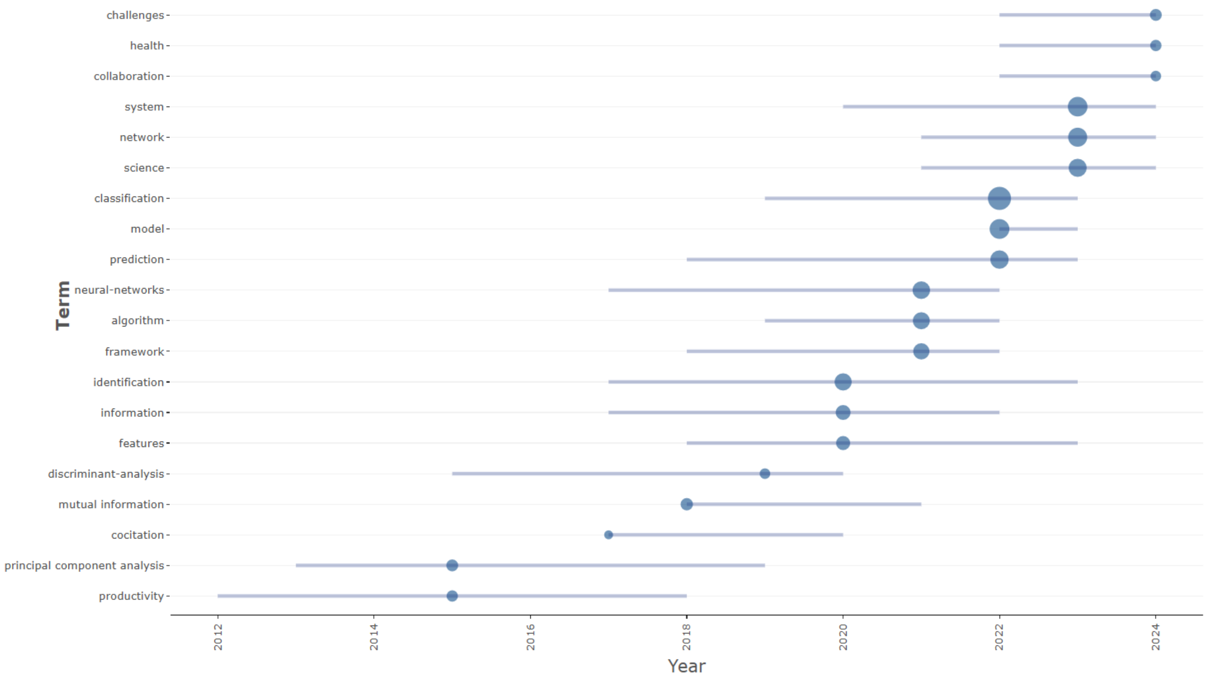

Figure 5. Temporal trends of keywords

Journal impact

Identifying the primary publication outlets is a critical step for comprehending the core communication channels and influential sources that disseminate research within this field. Table 2, which presents the top journals ranked according to various impact metrics (H-index, G-index, M-index, Total Citations-TC, Number of Publications-NP), reveals a varied portfolio of journals, mirroring the interdisciplinary nature of research on AI for knowledge discovery in innovation research. Leading the list is *Pattern Recognition*, with standout metrics (H-index of 27, 2410 total citations, and 86 publications). Other key venues include journals specialising in applied AI (*Expert Systems with Applications*), alongside prominent titles in specific engineering and application areas relevant to innovation dynamics (*Transportation Research Part C: Emerging Technologies*). Of particular significance is the inclusion of *Scientometrics* journal among these top-ranked sources. As a cornerstone publication in the SoS field, its presence strongly indicates that scholarly work applying AI to analyse scientific, technological, and innovation phenomena is not only published, but also holds significant recognition within the very community dedicated to studying science itself.

Source journal	H-index	G-index	M-index	TC	NP
Pattern Recognition	27	46	3.375	2410	86
Transportation Research Part C: Emerging Technologies	18	22	1.385	1673	22
Scientometrics	13	20	0.867	557	20
IEEE Access	12	22	1.333	515	35
IEEE *Transactions on Geoscience and Remote Sensing*	11	20	0.917	421	27
IEEE *Transactions on Neural Networks and Learning Systems*	10	19	0.833	388	19
Expert Systems with Applications	9	10	0.563	535	10
Pattern Recognition Letters	9	21	0.692	735	21
International Journal of Pattern Recognition and Artificial Intelligence	7	10	0.5	170	55
ACM *Transactions on Knowledge Discovery from Data*	6	9	1.2	85	18

Table 2. Top journals by impact metrics

The cumulative publication trends over time, shown in figure 6, illustrate the shifting dominance and growth trajectories of key journals. In recent years, several journals have shown rapid increases in their publication volume within this domain. The most notable growth is seen in *Pattern Recognition* (blue line), which exhibits an exceptionally steep upward curve, solidifying its position as the primary venue. Following this trend, the *International Journal of Pattern Recognition and Artificial Intelligence* (green line) and *IEEE Access* (red line) also demonstrate significant and sustained growth, establishing themselves as major contributors. Furthermore, *IEEE Transactions on Geoscience and Remote Sensing* (brown line) and *Transportation Research Part C: Emerging Technologies* (pink line) show considerable acceleration in their publication rates, indicating their increasing importance as outlets for this research. This visualisation effectively highlights a competitive and expanding publication landscape, with a clear recent surge across several leading journals.

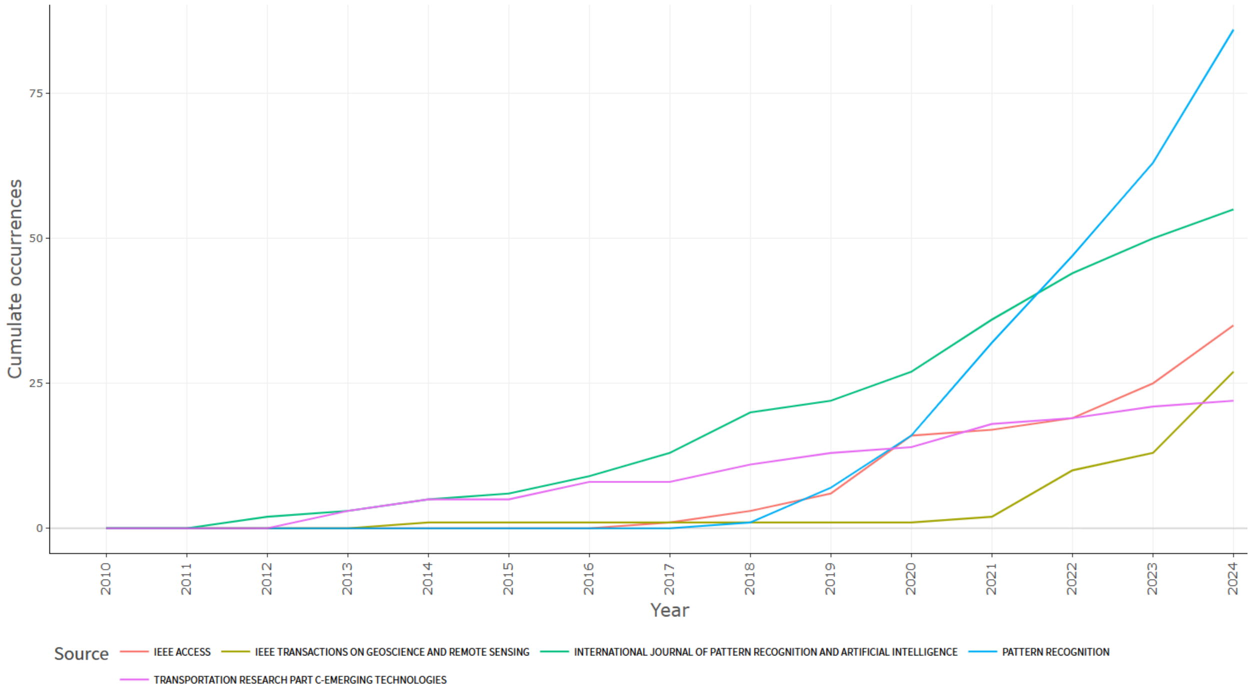

Figure 6. Cumulative journal publications over time

Documents

Table 3 provides crucial insights into the foundational works and highly impactful studies that have significantly shaped the intellectual contours of research on AI for knowledge discovery in innovation research. These publications, representing key reference points that have garnered substantial scholarly attention, are led by Wei et al. (2018) in *Renewable and Sustainable Energy Reviews* (534 total citations), followed by influential contributions from Mao et al. (2018) in *IEEE Communications Surveys and Tutorials* and Chavarriaga et al. (2013) in *Pattern Recognition Letters*. The diverse range of publication venues vividly underscores the field's interdisciplinary nature, with impactful papers originating from core AI and pattern recognition journals (e.g., *Pattern Recognition Letters*, *Swarm And Evolutionary Computation*, *Pattern Recognition*) alongside journals focusing on specific application domains highly relevant to innovation systems, such as energy, communications, transportation (*Transportation Research Part C: Emerging Technologies*), urban studies (*Cities*), and engineering (*Structural Health Monitoring*). Spanning publication years from 2013 to 2023, these highly cited works encompass both established foundational research and more recent breakthroughs. Metrics, like *TC per year* and *normalised TC*, further highlight the rapid and significant uptake of recent works, such as Malekloo et al. (2021) and Yuan et al. (2023). Collectively, these highly cited documents illuminate the key theoretical frameworks, methodological approaches, and application successes defining the intellectual core of this rapidly evolving field.

Paper	Source journal	Total citations	TC per year	Normalised TC
Wei et al. (2018)	*Renewable and Sustainable Energy Reviews*	534	66.75	9.43
Mao et al. (2018)	*IEEE Communications Surveys and Tutorials*	509	63.63	8.99
Chavarriaga et al. (2013)	*Pattern Recognition Letters*	483	37.15	3.86
Malekloo et al. (2021)	*Structural Health Monitoring*	295	73.75	11.44
Yuan et al. (2023)	*Pattern Recognition*	291	97.00	17.24
Nguyen et al. (2020)	*Swarm And Evolutionary Computation*	275	45.83	9.22
Javed et al. (2022)	*Cities*	265	66.25	10.28
Liu et al. (2018)	*International Journal of Computer Vision*	257	36.71	6.74
Gu et al. (2016)	*Transportation Research Part C: Emerging Technologies*	244	24.40	5.15
Li et al. (2013)	*Transportation Research Part C: Emerging Technologies*	243	18.69	1.94

Table 3. Top globally cited documents

Affiliations

Analysis of author affiliations provides a geographical and institutional perspective, identifying the leading organisations that are most actively contributing to research in this domain. Table 4, presenting the top affiliations by publication output, indicates a strong global presence in this study area, with a particularly notable concentration of highly productive institutions from China among the top ranks within this specific dataset. Universities and research bodies such as the Chinese Academy of Sciences (104 articles), Shenzhen University (sixty articles), Xidian University (forty articles), and Xi'an Jiaotong University (thirty-six articles), among others, demonstrate substantial scholarly output. This pattern points towards significant institutional focus, strategic investment, and concentrated expertise, dedicated to exploring the complex interplay between AI, knowledge discovery methodologies, and their application within the context of innovation research across these organisations. Understanding the geographical distribution and institutional strengths is thus vital for identifying potential research hubs and comprehending the global dynamics of knowledge production in this rapidly advancing interdisciplinary domain.

Affiliation	Articles
Chinese Academy of Sciences	104
Shenzhen University	60
Xidian University	40
Xi'an Jiaotong University	36
Shanghai Jiao Tong University	34
Chongqing University	31
Northwestern Polytechnical University	30
Egyptian Knowledge Bank	28
Tongji University	27
University of Chinese Academy of Sciences	27
Shandong University	23
Wuhan University	22

Central South University	21
Southern University of Science and Technology	21
Hunan University	20
Xiamen University	20
Nanchang Hangkong University	19
Tsinghua University	19
Sichuan University	18
Zhejiang University	18

Table 4. Top affiliations by publication output

Countries

Complementing the analysis of affiliations, table 5 maps the top countries, based on their total publication output, offering a broader geographical perspective on the global distribution of research productivity in this field. Consistent with the institutional-level data, China emerges as the leading country by a considerable margin in terms of publication frequency (2,172), followed by the USA (214) and South Korea (ninety-one). Other nations like India (seventy-eight), the UK (seventy-eight), Australia (sixty-eight), and Canada (fifty-four) also appear prominently in the rankings, confirming that research at this intersection is indeed a widely pursued international endeavour. It should be noted that the frequency metric represents the total number of appearances of a country in the affiliation lists of all documents; therefore, a single internationally co-authored paper will contribute to the frequency count of each participating country. The observed dominance of certain countries underscores the strategic importance assigned to leveraging AI and data-driven approaches for understanding and fostering innovation within these nations' research and development ecosystems. An exploration of the international collaboration patterns among these leading countries, a task reserved for the subsequent Network analysis section, will further illuminate the cross-border dynamics and collaborative structures that are actively shaping this evolving research landscape on a global scale.

Country	Frequency
China	2172
USA	214
South Korea	91
India	78
UK	78
Australia	68
Canada	54
Spain	45
Egypt	38
Germany	33
Italy	32
Brazil	30
Finland	29
France	27
Japan	26
Pakistan	26
Iran	25
Singapore	23
Malaysia	20
Vietnam	17

Table 5. Top countries by publication output

Network analysis

Co-occurrence network

The keyword co-occurrence network, visualised in figure 7, provides a structural map of the research themes animating the field of AI for knowledge discovery in innovation research. Generated using VOSviewer, this network elucidates the relationships between key concepts, based on their co-occurrence in the literature, with node size indicating frequency and colours representing distinct thematic clusters. Four primary clusters emerge, revealing the intellectual organisation of the domain. The largest cluster (red), situated on the left, appears to represent the core conceptual foundations, application contexts, and meta-analytical perspectives. It prominently features overarching terms like *artificial intelligence* and *machine learning* alongside innovation-centred keywords such as *innovation*, *technology*, *emerging technologies*, *patent analysis*, and *trends*. Crucially, this cluster also encompasses terms related to evaluation (*performance*, *impact*) and the methodologies used for analysing the scientific landscape itself, including *bibliometrics*, *scientometrics*, *text mining*, and *natural language processing*, directly linking the research to science of science practices.

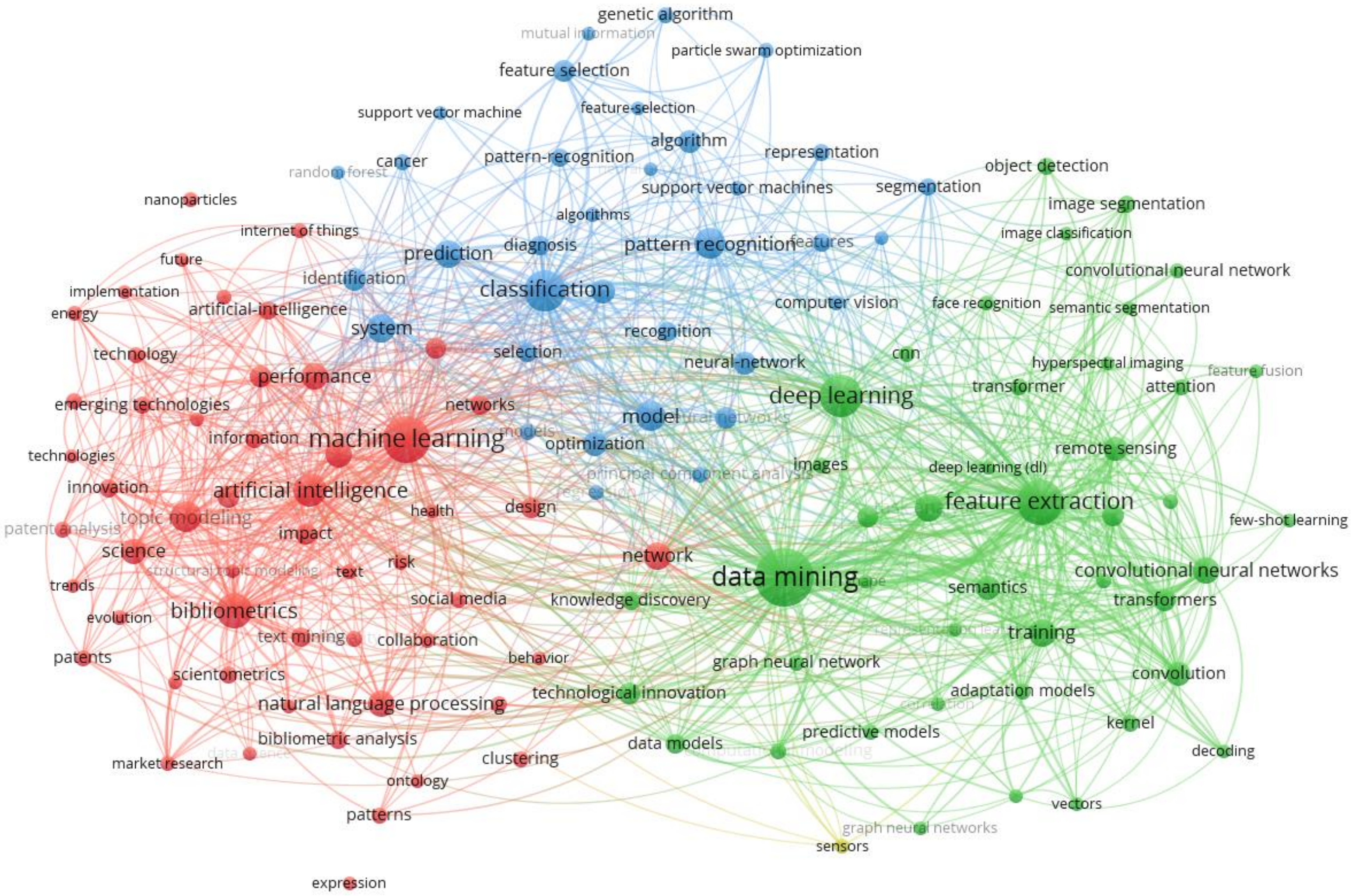

Figure 7. Keyword co-occurrence network

Adjacent to this, the blue cluster focuses on classical machine learning tasks and algorithms, featuring keywords like *classification*, *prediction*, *pattern recognition*, *feature selection*, *support vector machine*, *algorithm*, and *diagnosis*, suggesting a focus on predictive modelling and established recognition techniques. The green cluster, dominating the right side, clearly pertains to advanced AI techniques, data processing, and applications often involving unstructured data. Central nodes like *data mining* and *deep learning* are surrounded by terms such as *feature extraction*, *convolutional neural networks*, *transformers*, *image segmentation*, *remote sensing*, and *training*, highlighting the adoption of sophisticated deep learning architectures. Finally, a smaller, less distinct yellow cluster is discernible at the bottom, centred upon *sensors* and potentially linked to adjacent terms like *predictive models* and *graph neural networks*, hinting at research focused on specific data sources, or emerging modelling approaches related to sensor data or complex network structures. The significant interconnections between these clusters underscore the field's interdisciplinary nature, bridging core AI methods, advanced techniques, innovation studies, and self-reflective scientific analysis.

The structure of this network is highly revealing of the field's interdisciplinary identity and its dual nature. The large red cluster's role as a central hub connecting core AI methods, innovation contexts, and meta-analytical techniques is particularly significant. It demonstrates that researchers are engaged in a dual mission: (1) applying AI as a practical tool to analyse innovation (evidenced by terms like *patent analysis* and *trends*), and simultaneously (2) using the methodologies of the Science of Science (*bibliometrics*, *scientometrics*) to study and map the application of AI within this very context. This reflexivity, using the science of science to understand its own evolving methods, is a core tenet of the AI for science of science agenda and explains why these seemingly disparate concepts co-occur so frequently. The distinct blue and green clusters, representing classical and advanced AI techniques respectively, illustrate a methodological continuum, and their strong links to the red cluster show how these different computational tools are operationalised to address the core innovation challenges. Ultimately, the network map suggests a field that is actively integrating AI into a self-aware analytical framework for understanding scientific and technological progress.

Co-authorship network

Delving into the social structure underpinning knowledge creation in this field, figure 8 illuminates the landscape of scholarly collaboration through an author co-authorship network. This visualisation reveals several distinct, relatively dense clusters of researchers who collaborate frequently amongst themselves, connected by sparser links indicating inter-group collaboration. Each cluster appears to represent a community focused on specific methodological or application areas within the broader theme. For instance, the green cluster, prominently featuring Wei Wang (Chinese Academy of Sciences), suggests a focus on deep learning security, adversarial attacks, and data protection techniques. Adjacent to this, the yellow cluster led by Zhihui Lai (Shenzhen University) seems dedicated to advanced feature extraction methodologies, particularly manifold learning and adaptive graph embedding for pattern recognition. The blue cluster, centred on Licheng Jiao (Xidian University), indicates research into autoencoder architectures and feature learning, with specific applications in remote sensing image analysis like SAR change detection.

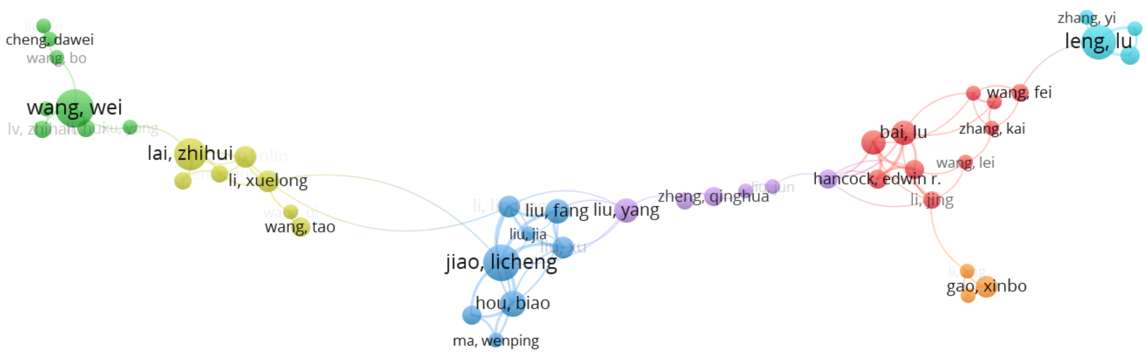

Figure 8. Author co-authorship network

Bridging towards the right, the purple cluster represented by Yang Liu (Dongguan Polytechnic) points towards applications in educational data mining and learning analytics, using deep learning for student risk prediction. Further right, the red cluster, with Lu Bai (Central University of Finance and Economics) as a key figure, delves into novel graph representation techniques, potentially drawing from quantum-inspired methods and entropy for graph analysis. Connected to this, the orange cluster involving Xinbo Gao (Chongqing University of Posts and Telecommunications) appears focused on multi-view clustering and graph learning, potentially utilising tensor methods. Lastly, the distinct indigo cluster led by Lu Leng (Nanchang Hangkong University) concentrates on biometric recognition, specifically palmprint analysis using specialised network architectures. The overall network structure, characterised by these specialised yet interconnected communities, highlights both focused expertise within groups and the essential cross-pollination of ideas across different research teams and sub-disciplines.

The formation of these distinct, yet interconnected, research communities is likely driven by a combination of institutional, thematic, and social factors. The high concentration of researchers from leading Chinese institutions (e.g., Chinese Academy of Sciences, Shenzhen University, Xidian University) within specific clusters suggests that geographical proximity, shared institutional resources, and established academic networks, play a significant role in fostering these collaborations. Furthermore, the highly specialised nature of each cluster, such as the yellow cluster's focus on manifold learning, or the blue cluster's dedication to remote sensing, points to collaborations coalescing upon specific, complex research problems that require sustained, in-depth exploration. These specialised groups may represent core research labs, or teams funded by specific national grants. This structure reveals a knowledge production model where deep expertise is cultivated within focused groups, while broader field progress is achieved through the cross-pollination of ideas and methods between these specialised hubs.

Co-citation network

The document co-citation network, presented in figure 9, illuminates the intellectual foundations upon which the research field of AI for knowledge discovery in innovation research is built. This map visualises which publications are frequently cited together within the analysed literature, revealing clusters of influential works that represent key knowledge pillars. Four distinct clusters emerge, signifying different, yet interconnected, streams of foundational influence. The large red cluster is anchored by seminal works in deep learning for computer vision, prominently featuring He et al. (2016) on Residual Networks (ResNet), alongside foundational contributions like LeCun et al.'s (1998) early convolutional neural networks (CNN) work and Deng et al.'s (2009) ImageNet dataset. Closely associated, the yellow cluster highlights other highly impactful vision architectures, potentially including works like Ronneberger et al. (2015) on U-Nets for segmentation or Hu et al. (2018) on attention mechanisms within convolutional neural networks, building upon the core CNN foundations.

Transitioning towards sequence-based models, the green cluster is centred on Vaswani et al. (2017), which introduced the revolutionary transformer architecture (*Attention Is All You Need*), fundamentally impacting natural language processing and beyond. This cluster also incorporates related works on sequence modelling and visualisation techniques like t-SNE (van der Maaten, 2008). Distinctly separate, yet connected through bridging citations, the blue cluster represents the domain of probabilistic topic modelling and bibliometric analysis tools. It is anchored by Blei et al. (2003) on Latent Dirichlet Allocation (LDA) and includes the pivotal work by van Eck and Waltman (2010) detailing the VOSviewer software, a tool fundamental to creating the very maps used in this study. The structure reveals a strong reliance on foundational deep learning literature (especially vision), alongside significant influence from breakthrough sequence models (Transformers) and key methods for text analysis (LDA) and science mapping (VOSviewer), reflecting the diverse methodological toolkit employed in the field.

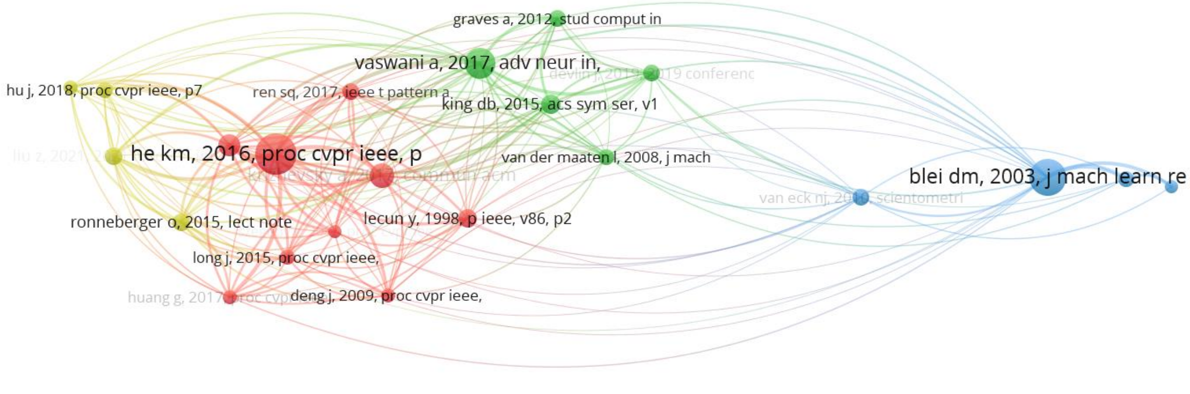

Figure 9. Document co-citation network

Bibliometric coupling

Mapping the current research fronts through shared intellectual foundations, figure 10 presents the document bibliometric coupling network. This visualisation connects documents within the dataset that cite similar references, thus highlighting clusters of contemporary research focused on related themes, or drawing upon a common knowledge base. Distinct clusters emerge, signifying active research areas citing similar foundational literature. For instance, a prominent grouping (red and blue) showcases recent applications of advanced deep learning, exemplified by Mao et al.'s (2018) work applying deep learning (DL) to wireless networks and Yuan et al.'s (2023) investigation into hybrid CNN-Transformer architectures for medical image segmentation. Adjacent research, represented by Wei et al.'s (2018) highly coupled work (indigo), delves into data-driven building energy analysis and prediction.

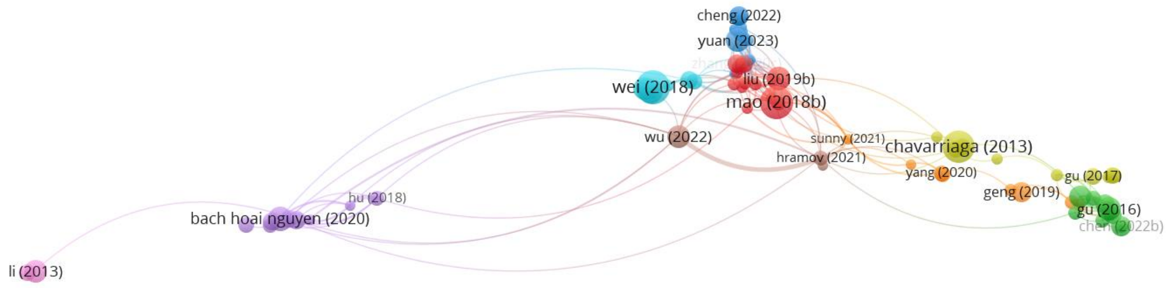

Figure 10. Document bibliometric coupling

Another significant constellation (yellow, green, orange) appears focused on leveraging real-world sensor or social data, addressing topics like human activity recognition benchmarking (Chavarriaga et al., 2013), traffic incident detection from social media (Gu et al., 2016), and array-based chemical sensing (Geng et al., 2019). More methodologically oriented research streams are also visible, such as the purple cluster surrounding Nguyen et al.'s (2020) work on swarm intelligence for feature selection, and the somewhat peripheral pink cluster linked to early work on traffic data imputation (Li et al., 2013). Furthermore, the distinct brown cluster (Wu et al., 2022) stands out, investigating the specialised area of transfer learning within brain-computer interfaces. The connections linking these varied clusters signify the shared foundational knowledge, likely common AI techniques or influential datasets, underpinning these diverse contemporary research directions. In short, this coupling map reveals a landscape characterised by distinct, but interconnected, research fronts, primarily driven by the application of sophisticated AI methodologies to diverse innovation-related challenges, spanning specific domain problems and methodological refinements.

Countries' collaboration world map

Visualising the global footprint of research partnerships, Figure 11 maps the international collaboration network, based on co-authorship between countries within the dataset. This geographical representation highlights the key national players and the intensity of their collaborative ties in the field of AI for knowledge discovery in innovation research. Dominating the landscape are China and the USA, depicted with darker shading and serving as major hubs with numerous connections radiating outwards. A particularly strong and dense set of links connects China and the USA, signifying a highly significant bilateral collaboration axis in this domain. Beyond this central dyad, China also demonstrates substantial collaborative activity with other prominent research nations, including the United Kingdom, Australia, Singapore, and Canada. Likewise, the USA maintains robust collaborative links not only with China, but also with various European countries and Canada. While research productivity is globally distributed, as shown in the descriptive analysis, this collaboration map underscores the central role played by China and the USA in driving international partnerships, forming the backbone of a globally interconnected network that facilitates knowledge exchange across continents.

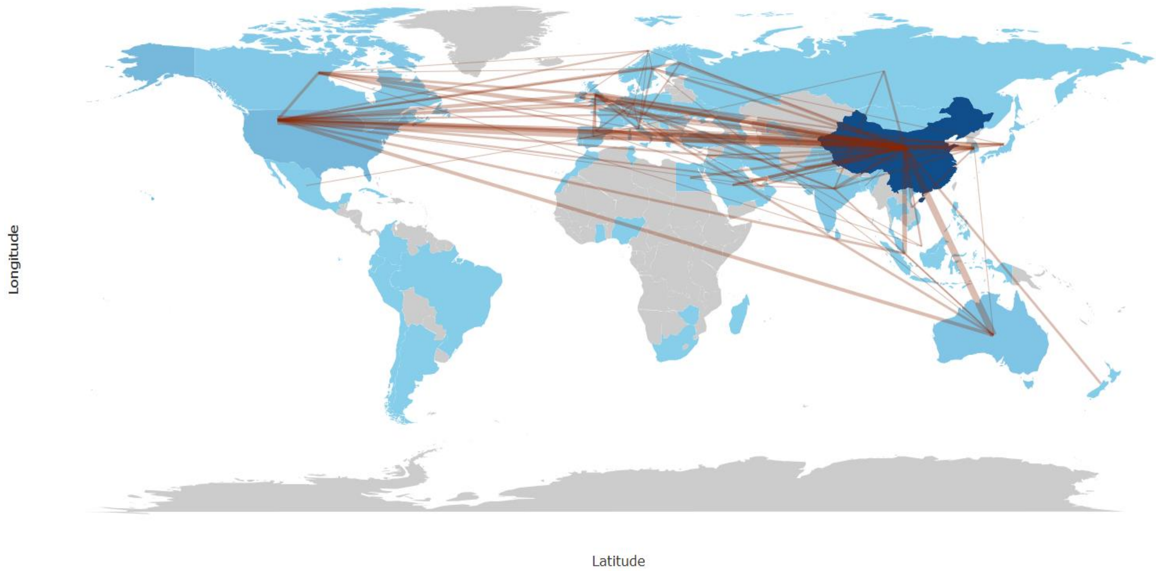

Figure 11. International collaboration network of countries

Discussion

Interpretation of the research landscape

The exponential growth observed in publications, marked by an impressive 37.52% annual growth rate and a low average document age, strongly signals the burgeoning importance and timeliness of applying AI to extract knowledge from innovation-related data. This rapid expansion likely stems from a confluence of factors: the maturation and accessibility of powerful AI techniques (deep learning, transformers), the increasing availability of large-scale digital innovation data (e.g., scientific literature, patent databases, online reports), and a growing recognition within academia and industry of the need for advanced analytical tools to navigate the complexity of modern innovation ecosystems (Secundo et al., 2025). The highly collaborative nature, evidenced by high co-authorship rates and significant international collaboration, underscores the field's inherent interdisciplinarity. Furthermore, the pronounced leadership in publication output, particularly from China, is not merely a reflection of research volume, but is also indicative of strong national strategic initiatives and targeted funding policies designed to establish advantages in critical technological areas like AI (Radu, 2021).

The thematic structure, visualised through the keyword co-occurrence network, reveals a multifaceted landscape. The red cluster acts as a crucial hub, integrating core AI concepts (*artificial intelligence*, *machine learning*) with innovation context (*innovation*, *patent analysis*, *emerging technologies*) and, significantly, meta-analytical methods (*bibliometrics*, *scientometrics*, *text mining*). This highlights that the field encompasses not only the application of AI to innovation, but also the use of scientific methods to study this very intersection, directly aligning with science of science (SoS) principles. The blue cluster represents the enduring utility of established machine learning tasks, such as classification and prediction, likely applied to structured innovation data or forecasting problems. The green cluster, centred on *data mining*, *deep learning*, and *feature extraction*, signifies the significant impact of advanced AI in handling complex, often unstructured innovation data (e.g., text, potentially images in patents or reports). The smaller yellow cluster upon *sensors* suggests nascent research exploring real-time innovation monitoring or specific data modalities. The temporal keyword trends further illustrate a potential shift from foundational techniques towards a more recent, intense focus on core AI applications (*classification*, *network*, *model*) and evaluating their effectiveness (*performance*, *impact*).

The intellectual foundations, mapped by the co-citation network, are revealing. The strong presence of foundational works in computer vision (He *et al.*, 2016) and natural language processing (NLP; Vaswani et al., 2017) underscores the importance of techniques developed for processing image and sequence data, likely applied to patent diagrams, product images, or the vast textual corpus of scientific articles and reports relevant to innovation. Critically, the co-citation of these AI cornerstones alongside key works in topic modelling (Blei et al., 2003) and science mapping software (van Eck & Waltman, 2010) provides compelling evidence for the integration of AI methods with established text analysis and scientometric tools. This unique blend forms the intellectual bedrock of the field. While the field builds on these strong foundations, its rapid evolution also presents complex dynamics. The observed decline in average citations for recent years, while primarily a function of a shorter citation window, may also be compounded by the rapid proliferation of niche topics and specialized sub-fields, a common phenomenon in fast-maturing scientific domains where impact takes time to diffuse across different research fronts (Sun & Latora, 2020). This diversification is also reflected in the specialised author clusters, which point to a vibrant, but increasingly segmented, social structure.

Implications and significance

This study's primary contribution lies in providing the first comprehensive, data-driven map of the research landscape at the intersection of AI, knowledge discovery, and innovation research, addressing a significant gap in the literature. By delineating the key themes, influential actors, intellectual structure, and collaborative patterns, this work offers a valuable navigational aid for researchers and a baseline for future studies tracking the field's evolution. Beyond this specific domain, the findings hold significant implications for the broader field of AI for SoS (science of science). On the one hand, this study empirically demonstrates how AI is actively enabling and transforming SoS and innovation studies. AI techniques, particularly deep learning and advanced natural language processing, are empowering researchers to move beyond traditional bibliometric indicators to analyse the content and context of innovation data (scientific text, patents, reports) at unprecedented scale and depth, facilitating more nuanced knowledge discovery about scientific and technological progress (Agrawal et al, 2024). The prominence of clusters related to deep learning and feature extraction attests to this capability.

On the other hand, and equally importantly, the findings highlight how science of science methodologies are being employed to understand and structure the very field where AI meets innovation analysis. The significant presence of keywords like *bibliometrics* and *scientometrics* within the core thematic cluster, the appearance of *Scientometrics* among top journals, and the citation of foundational scientometric software reveal a meta-analytical perspective inherent within the field. Researchers are not just using AI for innovation studies; they are also using science mapping and analysis techniques (often enhanced by AI itself, e.g., natural language processing for text mining) to study how this application field is developing. This reflexivity is central to the AI for SoS concept. This convergence, however, also surfaces new science of science challenges: how can we rigorously evaluate the actual impact of AI-driven discoveries on innovation outcomes beyond citation metrics? How do we address potential biases embedded in AI algorithms when used for evaluating research or identifying trends (Moon and Ahn, 2025)? This study lays the groundwork for exploring these critical questions. Practically, the map provided can guide practitioners in identifying relevant AI tools for technology intelligence, market analysis, and R&D management, while informing policymakers about global research strengths and potential areas for strategic investment in AI-driven innovation analytics.

Limitations and future research directions

Whilst this study offers valuable insights, certain limitations must be acknowledged, which in turn suggest avenues for future research. Firstly, this study's reliance solely on the WoS Core Collection, while ensuring data quality and consistency, is a primary limitation. It inevitably excludes potentially relevant publications indexed in other major databases, most notably Scopus. Although using a single, high-quality database was a deliberate methodological choice, to ensure data integrity for this initial mapping, we acknowledge the reviewer's valuable suggestion that future research could achieve greater comprehensiveness. Tools such as the Bibliometrix package are specifically designed to merge and deduplicate data from multiple sources like WoS and Scopus. Therefore, a promising future direction is to conduct a comparative or integrated analysis using both databases to capture a more complete global landscape and validate the trends identified here. Secondly, the keyword-based search strategy, while carefully constructed, might not capture all relevant literature, particularly studies that describe applications without explicitly using the chosen set of *knowledge discovery* or *innovation research* terms. Refining search strategies, or employing citation expansion techniques, could mitigate this. Thirdly, bibliometric analysis primarily reveals structural patterns and trends; it offers limited insight into the qualitative nuances, theoretical depth, or practical validity of the research content itself.

Consequently, future research should move towards integrating these quantitative mapping results with qualitative approaches. In-depth content analysis of publications within key thematic clusters or research frontiers could provide richer understanding of the specific problems being addressed and the AI methods employed. Longitudinal thematic evolution analysis (using tools like Bibliometrix's built-in functions) could offer a more dynamic picture than the keyword trend snapshot. Furthermore, comparative studies evaluating the performance of different AI techniques (e.g., various deep learning models vs. traditional machine learning) for specific innovation knowledge discovery tasks (e.g., technology forecasting accuracy, patent novelty detection) are needed. Crucially, research aligned with the AI for SoS agenda should focus on developing new frameworks and metrics to assess the tangible impact of AI on the efficiency, direction, and equity of scientific discovery and innovation processes. Investigating the ethical dimensions, such as algorithmic bias in trend detection, or fairness in AI-assisted research evaluation, is paramount. Fostering deeper interdisciplinary collaboration between AI experts, innovation scholars, economists, and policy analysts will be essential to fully realise the potential and address the challenges at the intersection of AI, knowledge discovery, and the science of innovation.

Conclusion

This bibliometric exploration has charted the vibrant and rapidly accelerating landscape where artificial intelligence serves as a catalyst for knowledge discovery within innovation research. More than a mere description of a burgeoning field, our analysis illuminates a profound, symbiotic relationship in which AI is simultaneously a transformative analytical tool and an emerging object of study within the science of science (SoS) framework itself. The field's trajectory is marked by exponential publication growth and a highly collaborative, international research environment, propelled by a concentrated cohort of institutions with significant strategic focus apparent in global research powerhouses such as China and the USA. Intellectually, the landscape is not monolithic but is organised upon a sophisticated triumvirate of core themes: foundational AI methodologies, their application in specific innovation contexts, and, most revealingly, the meta-analytical methods of the science of science (SoS), confirming the field's dual identity as both a user and an object of scientific self-study. This intellectual architecture is mirrored in its social structure, characterised at the macro level by a robust Sino-American partnership and at the micro level by distinct, densely connected author clusters dedicated to specialised problems. The integration of AI into innovation studies therefore represents far more than a methodological update; it signifies a fundamental shift in the cognitive infrastructure of the field, moving from descriptive analysis toward predictive modelling and potentially automated discovery. Realising this future responsibly, however, demands a clear-eyed approach to developing robust frameworks to evaluate AI's true impact, mitigating the inherent risks of algorithmic bias, and fostering genuine, synergistic collaboration. The thoughtful stewardship of this powerful convergence holds the key, not only to advancing innovation research, but also to unlocking new frontiers of understanding in an increasingly complex world.

About the author

Yanyi Wu is with the School of Public Affairs and the Institute of China's Science, Technology and Education Policy, Zhejiang University, Hangzhou, China. His research interests include the governance of emerging technologies. He can be contacted at yanyi.wu@hotmail.com

Reference

Agrawal, A., McHale, J., & Oettl, A. (2024). Artificial intelligence and scientific discovery: a model of prioritized search. *Research Policy*, 53(5), 104989. https://doi.org/10.1016/j.respol.2024.104989

Aria, M., & Cuccurullo, C. (2017). Bibliometrix: an R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Ayinaddis, S. G. (2025). Artificial intelligence adoption dynamics and knowledge in SMEs and large firms: a systematic review and bibliometric analysis. *Journal of Innovation & Knowledge*, 10(3), 100682. https://doi.org/10.1016/j.jik.2025.100682

Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022. https://doi.org/10.5555/944919.944937

Bogers, M., Chesbrough, H., & Moedas, C. (2018). Open innovation: Research, practices, and policies. *California Management Review*, 60(2), 5–16. https://doi.org/10.1177/0008125617745086

Callon, M., Courtial, J. P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: an introduction to co-word analysis. *Social Science Information*, 22(2), 191–235. https://doi.org/10.1177/053901883022002003

Chavarriaga, R., Sagha, H., Calatroni, A., Digumarti, S. T., Tröster, G., Millán, J. del R., & Roggen, D. (2013). The opportunity challenge: A benchmark database for on-body sensor-based activity recognition. *Pattern Recognition Letters*, 34(15), 2033–2042. https://doi.org/10.1016/j.patrec.2012.12.014

Chellappa, R., Theodoridis, S., & van Schaik, A. (2021). Advances in machine learning and deep neural networks. *Proceedings of the IEEE*, 109(5), 607–611. https://doi.org/10.1109/jproc.2021.3072172

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. 2009 *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 248–255). https://doi.org/10.1109/cvpr.2009.5206848

Dessimoz, C., & Thomas, P. D. (2024). AI and the democratization of knowledge. *Scientific Data*, 11, 268. https://doi.org/10.1038/s41597-024-03099-1

Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., Vespignani, A., Waltman, L., Wang, D., & Barabási, A.-L. (2018). Science of science. *Science*, 359(6379), eaao0185. https://doi.org/10.1126/science.aao0185

Gama, F., & Magistretti, S. (2025). Artificial intelligence in innovation management: A review of innovation capabilities and a taxonomy of AI applications. *Journal of Product Innovation Management*, 42(1), 76–111. https://doi.org/10.1111/jpim.12698

Geng, Y., Peveler, W. J., & Rotello, V. M. (2019). Array-based "Chemical Nose" sensing in diagnostics and drug discovery. *Angewandte Chemie*, 58(16), 5190–5200. https://doi.org/10.1002/anie.201809607

Gu, Y., Qian, Z. (S.) & Chen, F. (2016). From twitter to detector: Real-time traffic incident detection using social media data. *Transportation Research Part C: Emerging Technologies*, 67, 321–342. https://doi.org/10.1016/j.trc.2016.02.011

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition* (CVPR), 770–778. https://doi.org/10.1109/cvpr.2016.90

Howoldt, D. (2024). Characterising innovation policy mixes in innovation systems. *Research Policy*, 53(2), 104902. https://doi.org/10.1016/j.respol.2023.104902

Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7132–7141). https://doi.org/10.1109/cvpr.2018.00745

Jarrahi, M. H., Askay, D., Eshraghi, A., & Smith, P. (2022). Artificial intelligence and knowledge management: A partnership between human and AI. *Business Horizons*, 66(1), 87–99. https://doi.org/10.1016/j.bushor.2022.03.002

Javed, A. R., Shahzad, F., Rehman, S. ur, Zikria, Y. B., Razzak, I., Jalil, Z., & Xu, G. (2022). Future smart cities requirements, emerging technologies, applications, challenges, and future aspects. *Cities*, 129(1), 103794. https://doi.org/10.1016/j.cities.2022.103794

Katz, J. Sylvan., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1–18. https://doi.org/10.1016/s0048-7333(96)00917-1

Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 14(1), 10–25. https://doi.org/10.1002/asi.5090140103

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2022). Natural language processing: State of the art, current trends and challenges. *Multimedia Tools and Applications*, 82(3), 3713–3744. https://doi.org/10.1007/s11042-022-13428-4

Lê, J. K., & Schmid, T. (2020). The practice of innovating research methods. *Organizational Research Methods*, 25(2), 308–336. https://doi.org/10.1177/1094428120935498

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324. https://doi.org/10.1109/5.726791

Li, L., Li, Y., & Li, Z. (2013). Efficient missing data imputing for traffic flow by considering temporal and spatial dependence. *Transportation Research Part C: Emerging Technologies*, 34, 108–120. https://doi.org/10.1016/j.trc.2013.05.008

Liu, L., Chen, J., Fieguth, P., Zhao, G., Chellappa, R., & Pietikäinen, M. (2018). From BoW to CNN: Two decades of texture representation for texture classification. *International Journal of Computer Vision*, 127(1), 74–109. https://doi.org/10.1007/s11263-018-1125-z

Malekloo, A., Ozer, E., AlHamaydeh, M., & Girolami, M. (2021). Machine learning and structural health monitoring overview with emerging technology and high-dimensional data source highlights. *Structural Health Monitoring*, 21(4), 147592172110368. https://doi.org/10.1177/14759217211036880

Mao, Q., Hu, F., & Hao, Q. (2018). Deep learning for intelligent wireless networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 20(4), 2595–2621. https://doi.org/10.1109/COMST.2018.2846401

Mariani, M. M., Machado, I., Magrelli, V., & Dwivedi, Y. K. (2023a). Artificial intelligence in innovation research: A systematic review, conceptual framework, and future research directions. *Technovation*, 122, 102623. https://doi.org/10.1016/j.technovation.2022.102623

Mariani, M. M., Machado, I. & Nambisan, S. (2023b). Types of innovation and artificial intelligence: a systematic quantitative literature review and research agenda. *Journal of Business Research*, 155, 113364https://doi.org/10.1016/j.jbusres.2022.113364

Marino, D., Gil Lafuente, J., & Tebala, D. (2023). Innovations and development of artificial intelligence in Europe: Some empirical evidences. *European Journal of Management and Business Economics*, 32(5), 620–636. https://doi.org/10.1108/ejmbe-03-2023-0085

Maruccia, Y., Solazzo, G., Del Vecchio, P., & Passiante, G. (2020). Evidence from network analysis application to innovation systems and quintuple helix. *Technological Forecasting and Social Change*, 161, 120306. https://doi.org/10.1016/j.techfore.2020.120306

Moon, D., & Ahn, S. (2025). Metrics and algorithms for identifying and mitigating bias in AI design: a counterfactual fairness approach. *IEEE Access*, 13, 59118–59129. https://doi.org/10.1109/access.2025.3556082

Nguyen, B. H., Xue, B., & Zhang, M. (2020). A survey on swarm intelligence approaches to feature selection in data mining. *Swarm and Evolutionary Computation*, 54, 100663. https://doi.org/10.1016/j.swevo.2020.100663

Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. *Policy and Society*, 40(2), 178–193. https://doi.org/10.1080/14494035.2021.1929728

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention: MICCAI 2015 (Lecture Notes in Computer Science, Vol. 9351; pp. 234–241). https://doi.org/10.1007/978-3-319-24574-4_28

Secundo, G., Spilotro, C., Gast, J., & Corvello, V. (2025). The transformative power of artificial intelligence within innovation ecosystems: A review and a conceptual framework. *Review of Managerial Science*, 19, 2697–2728. https://doi.org/10.1007/s11846-024-00828-z

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4), 265–269. https://doi.org/10.1002/asi.4630240406

Sun, Y., & Latora, V. (2020). The evolution of knowledge within and across fields in modern physics. *Scientific Reports*, 10, 12097. https://doi.org/10.1038/s41598-020-68774-w

van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(86), 2579–2605. https://www.jmlr.org/papers/v9/vandermaaten08a.html

(Archived in the Internet Archive at https://web.archive.org/web/20250124104847/https://jmlr.org/papers/v9/vandermaaten08a.html

van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In NIPS'17: *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 6000–6010). https://dl.acm.org/doi/10.5555/3295222.3295349

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365–391. https://doi.org/10.1016/j.joi.2016.02.007

Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., Chandak, P., Liu, S., Van Katwyk, P., Deac, A., Anandkumar, A., Bergen, K., Gomes, C. P., Ho, S., Kohli, P., Lasenby, J., Leskovec, J., Liu, T.-Y., Manrai, A., & Marks, D. (2023). Scientific discovery in the age of artificial intelligence. *Nature*, 620, 47–60. https://doi.org/10.1038/s41586-023-06221-2

Wei, Y., Zhang, X., Shi, Y., Xia, L., Pan, S., Wu, J., Han, M., & Zhao, X. (2018). A review of data-driven approaches for prediction and classification of building energy consumption. *Renewable and Sustainable Energy Reviews*, 82, 1027–1047. https://doi.org/10.1016/j.rser.2017.09.108

Wu, D., Xu, Y., & Lu, B.-L. (2022). Transfer learning for EEG-based brain–computer interfaces: A review of progress made since 2016. *IEEE Transactions on Cognitive and Developmental Systems*, 14(1), 4–19. https://doi.org/10.1109/TCDS.2020.3007453

Yuan, F., Zhang, Z., & Fang, Z. (2023). An effective CNN and transformer complementary network for medical image segmentation. *Pattern Recognition*, 136, 109228. https://doi.org/10.1016/j.patcog.2022.109228