



25 years of knowledge and information systems: a bibliometric retrospective

Luciano Barcellos-Paula^{1,2} · José M. Merigó³ · Ana M. Gil-Lafuente⁴

Received: 12 December 2025 / Revised: 18 February 2026 / Accepted: 12 March 2026
© The Author(s) 2026

Abstract

In 2025, *Knowledge and Information Systems* (KAIS) celebrates its 25th anniversary. To mark this occasion, this study presents a comprehensive bibliometric analysis covering the period 1999–2024, examining publication and citation patterns, influential authors, institutions, countries, and evolving research themes. Drawing on data from the Web of Science (WoS) Core Collection, the analysis integrates classical bibliometric indicators with science mapping techniques, including co-citation, bibliographic coupling, and keyword co-occurrence, supported by VOSviewer and Bibliometrix tools. The findings reveal that KAIS has published 2,385 articles, which have collectively received 54,319 citations, demonstrating sustained global influence. The journal's intellectual core remains grounded in data mining, classification, clustering, and machine learning, while emerging themes—such as deep learning, graph neural networks, natural language processing, and concept drift—reflect its adaptation to contemporary developments in artificial intelligence. China and the United States lead in both productivity and impact, illustrating the globalization of the KAIS research community. This study offers insights into the evolution of knowledge and information systems as an interdisciplinary field, emphasizing KAIS's dual role in preserving methodological rigor and promoting innovation. Overall, the journal stands as a benchmark for studying the dynamics of data-driven, intelligent, and responsible knowledge discovery.

✉ José M. Merigó
Jose.Merigo@uts.edu.au
Luciano Barcellos-Paula
lbarcellosdepaula@pucp.edu.pe
Ana M. Gil-Lafuente
amgil@ub.edu

¹ Pontificia Universidad Católica del Perú, Av. Universitaria 1801, San Miguel, Lima 15088, Perú

² CENTRUM Católica Graduate Business School, Jirón Daniel Alomía Robles 125, Urbanización Los Álamos de Monterrico, Santiago de Surco, Lima 15023, Perú

³ School of Computer Science, Faculty of Engineering and Information Technology, University of Technology Sydney, 81 Broadway, Ultimo, NSW 2007, Australia

⁴ Department of Business Administration, University of Barcelona, Av. Diagonal 690, 08034 Barcelona, Spain

Keywords Knowledge · Information systems · Bibliometrics · Web of science · VOSviewer · Bibliometrix

1 Introduction

The journal *Knowledge and Information Systems* (KAIS) serves as a prominent scholarly platform that facilitates the exchange of expertise and the dissemination of advances in knowledge systems and advanced information systems. Covering a broad spectrum of topics—including knowledge and data engineering, decision support, soft computing, and uncertainty management—KAIS has published a total of 2,385 articles between 1999 and 2024, which have collectively received 54,319 citations. This body of work reflects the contributions of researchers affiliated with leading institutions across five continents, underscoring the journal's dual commitment to foundational research and practical applications.

The academic reputation of KAIS is further evidenced by its citation metrics. According to the Web of Science (WoS) Journal Citation Reports 2024 [18], the journal is indexed in the Science Citation Index Expanded and has achieved an impact factor of 3.1. This positions KAIS 95th among 204 journals in the WoS category “Computer Science, Artificial Intelligence”, and 111th out of 258 journals in “Computer Science, Information Systems”. Additionally, in the 2024 Scopus rankings, KAIS attained a CiteScore of 5.7, ranking 141st among 474 journals in the “Information Systems” category.

Established in 1999, KAIS is currently led by Editor-in-Chief Professor Xindong Wu of Hefei University of Technology, China. In celebration of the journal's 25th anniversary, this study conducts a comprehensive bibliometric analysis of publications indexed in the WoS Core Collection [18] up to 2024. The analysis aims to examine publication and citation trends, identify influential works, and highlight the most productive authors, institutions, and countries or regions. Furthermore, it provides a comparative evaluation with other leading journals in the field, employing tools such as VOSviewer and bibliometric techniques including bibliographic coupling, co-citation analysis, and keyword co-occurrence mapping.

Bibliometric analysis serves an essential purpose in commemorating journal anniversaries, as it offers a systematic examination of publication patterns, reveals key contributors, and highlights emerging research domains. Such evaluations not only celebrate a journal's legacy but also guide its future trajectory by identifying areas of influence and growth. For instance, Cobo et al. [21] conducted a bibliometric review of *Knowledge-Based Systems* over its first 25 years, while La Paz et al. [77] summarized major research trends in the 25th anniversary edition of the *Information Systems Journal*. Similarly, to mark its 40th anniversary, *Fuzzy Sets and Systems* presented a bibliometric overview identifying the most productive authors, institutions, and countries, along with the structural evolution of publications and citations [60].

Despite the growing interest in bibliometric retrospectives of major information systems journals, no comprehensive analysis has yet been conducted for KAIS. Considering its 25 years of continuous publication and growing influence within the scientific community, a systematic evaluation of its intellectual structure, publication dynamics, and global research impact is both timely and necessary. To address this gap, the present study adopts a comprehensive bibliometric approach based on publications indexed in the WoS Core Collection. It combines performance analysis—examining publication output, citation impact, and leading contributors—with science mapping techniques, including co-citation, bibliographic coupling, and keyword co-occurrence analyses, supported by visualization tools

such as VOSviewer. This study therefore seeks to fill this gap by offering an evidence-based overview of KAIS's scholarly development, providing insights that can inform future editorial strategies and enhance understanding of its contribution to the knowledge and information systems discipline.

Beyond describing internal bibliometric patterns, this study positions KAIS within the broader evolution of knowledge and information systems research worldwide. At a global level, the field has progressed through three broad phases: (i) symbolic and rule-based knowledge systems in the 1990s, (ii) statistical learning and large-scale data mining in the 2000s, and (iii) data-centric and representation-driven artificial intelligence since the mid-2010s. The longitudinal evidence presented in this study enables a comparison between these global transformations and the thematic trajectory of KAIS.

The remainder of this article is structured as follows. Section 2 reviews the theoretical foundations and methodological evolution of bibliometrics. Section 3 describes the research methodology and data collection process. Section 4 presents the results of the bibliometric analysis, examining publication and citation patterns, highlighting influential articles, identifying citing works, and profiling key contributors, including authors, institutions, and countries. Section 5 explores the intellectual structure of the field through co-citation, bibliographic coupling, and co-occurrence network analyses. Finally, Sect. 6 concludes the study with key insights and directions for future research.

2 Literature review

The continuous growth of scientific production and the increasing complexity of interdisciplinary research have made bibliometric analysis a fundamental tool for understanding knowledge creation and dissemination. In recent years, bibliometrics has evolved from a descriptive practice into a data-driven discipline that integrates artificial intelligence, network science, and computational modeling. This section reviews the theoretical foundations and methodological evolution of bibliometrics, emphasizing its application in computer science and information systems (CS/IS) and its role in assessing individual journals. By synthesizing seminal works and recent advances, it aims to provide a comprehensive overview of how bibliometric methods contribute to mapping scientific progress and guiding editorial and research strategies.

2.1 Bibliometrics

Bibliometrics is a subfield of information and library science that applies quantitative and statistical methods to study patterns of scholarly communication. It was first defined by Pritchard [79] as the application of mathematical and statistical techniques to books and other communication media. Broadus [14] articulated a clear definition of bibliometrics as a domain of study that employs quantitative methods to analyse bibliographic data, while Rousseau [81] characterised it as the comprehensive evaluation of all elements connected to the publishing and readership of books and documents.

The origins of bibliometrics trace back to Garfield [34], whose concept of citation indexing transformed how scientific influence could be measured. The creation of the Science Citation Index (SCI) in the 1960s established citation analysis as a core instrument for evaluating scholarly communication [9]. Garfield's innovation provided an empirical basis for tracing

intellectual influence across disciplines, fundamentally shaping modern science evaluation [90].

The primary goal of bibliometrics is to collect and analyse bibliographic data—authorship, institutional affiliations, citations, and keywords—to uncover research trends, collaboration patterns, and intellectual structures [27]. Technological advances and the rise of digital databases have amplified the field's reach, facilitating large-scale analyses that reveal scientific evolution over time [7]. While early studies focused on publication and citation counts, more recent works employ science mapping and network visualization to expose thematic evolution and cognitive structures within disciplines [25].

As the field matured, bibliometrics became increasingly intertwined with data science, network theory, and visualization techniques, producing models that combine descriptive, relational, and predictive analysis [36]. Emerging subfields such as altmetrics and webometrics have extended its methods to online and social media contexts, reflecting the diversification of scholarly communication in the digital age [56]. Today, bibliometrics supports the evaluation of research performance, informs science policy, and assists institutions in identifying emerging research fronts.

2.2 Bibliometric studies in computer science and information systems

Bibliometric studies in computer science and information systems (CS/IS) have expanded substantially with the acceleration of digital transformation and interdisciplinary research. Early analyses focused on mapping the intellectual structure of computational intelligence and decision-making using citation and co-word analysis [19, 94]. The advent of visualization tools such as VOSviewer [93] and SciMAT [20] revolutionized the field by enabling the visual representation of collaboration networks, thematic clusters, and longitudinal topic evolution.

A significant stream of research has examined fuzzy systems and intelligent computing—core areas in CS/IS. Studies have traced the evolution of fuzzy set theory across journals, institutions, and nations [32] and analysed milestone journals such as *Fuzzy Sets and Systems* [60], *Information Sciences* [69], and the *International Journal of Fuzzy Systems* [104]. These analyses reveal a progression from traditional fuzzy logic toward hybrid paradigms integrating machine learning, optimization, and soft computing.

Beyond fuzzy research, bibliometric analyses have explored fast-evolving subfields such as deep learning [64], blockchain [38], artificial intelligence [29], and big data [99]. These studies provide macroscopic insight into how computational paradigms evolve, merge, and diffuse across disciplines. Other research has ranked institutions and countries by productivity and impact, revealing global trends in computer science research [107].

Recent contributions published in *Knowledge and Information Systems* (KAIS) highlight a growing convergence between bibliometrics, artificial intelligence, and data-driven evaluation. Koca and Çiftçi [55] mapped the intellectual structure linking big data and cybersecurity, identifying clusters related to data privacy, intrusion detection, and IoT integration. Mukhtiar et al. [73] proposed a contra-harmonic mean-based index that outperforms traditional *h*-type metrics for researcher assessment, while Mustafa et al. [74] introduced interpretable decision-rule models that combine machine learning and bibliometric data to identify high-impact scholars.

Other KAIS studies have applied bibliometric methods to specialized technological domains. Narayanan and Manimaran [75] analysed trends in geographic information systems (GIS), demonstrating a shift toward semantic modeling and decision-support systems. Tzika-Kostopoulou et al. [92] conducted a large-scale review of big data in transportation,

revealing dominant themes such as urban mobility and smart cities, and Yazdanjue et al. [102] offered the first comprehensive bibliometric study of social network anonymization, proposing a taxonomy of anonymization techniques. Together, these works illustrate how bibliometric analysis in CS/IS has evolved from static performance measurement toward integrated, AI-driven, and domain-specific frameworks that enhance transparency and interpretability in scientific evaluation.

2.3 Bibliometric studies for individual journals

Bibliometric analyses of individual journals have become a hallmark of academic reflection, particularly during milestone anniversaries. Such studies offer a systematic view of publication trends, citation patterns, and thematic shifts, serving both to celebrate a journal's history and to guide its strategic development. Prestigious outlets such as *Nature* and *The Lancet* have used anniversary editions to explore their scientific and societal influence [37, 72].

This tradition originated in the 1980s, when major economics and finance outlets began to evaluate their scholarly evolution [39]. The *Journal of Finance* presented one of the earliest internal citation reviews to document its research trajectory [40], followed by retrospectives in the *Journal of Financial Economics* [83] and the *American Economic Review* [6]. Other examples include retrospectives from the *Journal of Environmental Economics and Management* [57], and the *Journal of Economic Psychology* [52].

In management and business, anniversary bibliometric reviews are common tools for assessing scholarly influence and editorial direction. The *Journal of International Business Studies* conducted a 25-year analysis emphasizing international collaboration [44], while the *Journal of Business Research* and the *European Management Journal* each published comprehensive retrospective evaluations highlighting citation impact and emerging themes [11, 26]. Similar efforts have been made by the *International Journal of Hospitality Management* [67], *The Leadership Quarterly* [33], *Journal of International Management* [58] and *Cross Cultural & Strategic Management* [59], each offering valuable insights into decades of scholarly development.

In engineering and computer science, anniversary bibliometrics have become equally significant. *Computers & Industrial Engineering* marked its 40th anniversary with a bibliometric analysis of author and keyword networks [17, 97] for the 50th anniversary of the *International Journal of Systems Science*. Calma et al. [16] presented a general overview of *Operations Research*, García-Merino et al. [31] of *Technovation*, Shukla et al. [86] of *Engineering Applications of Artificial Intelligence*, and Yu et al. [105] of *Fuzzy Optimization and Decision Making*.

Within computer and information sciences, these traditions are deeply embedded. The *International Journal of Intelligent Systems* celebrated its 30th anniversary with a detailed review of author influence and thematic evolution [68], while *Information Sciences* marked its 50th anniversary through a large-scale mapping of topic evolution and citation interconnections [69]. Similar work has been developed for *Applied Intelligence* [66], *IEEE Transactions on Fuzzy Systems* [103] and *International Journal of Machine Learning and Cybernetics* [100].

These analyses demonstrate how journals employ bibliometrics not only to commemorate their legacy but also to guide future editorial policies and strengthen disciplinary identity.

Overall, the evolution of bibliometric research—from its conceptual foundations to modern, domain-specific applications—reflects the growing complexity and analytical maturity

of the field. In computer science and information systems, bibliometrics now bridges classical citation analysis with artificial intelligence and data science, enabling deeper insight into research performance, thematic evolution, and collaboration networks. Concurrently, journal-specific studies, particularly those marking anniversaries, serve as institutional mirrors that both celebrate intellectual heritage and inform future strategies. Together, these developments underscore bibliometrics' central role in understanding, managing, and forecasting the dynamics of contemporary scientific knowledge production.

3 Methods

This study employs key bibliometric indicators, including the total number of publications, the citation-per-paper ratio, the h -index [41, 70], and citation thresholds. The h -index measures both productivity and influence, indicating the number of papers (h) that have received at least h citations each. Although it effectively combines productivity and impact, it also presents limitations, particularly regarding outliers [3]. Citation thresholds, in turn, denote the number of documents reaching specific citation milestones, such as ten or one hundred citations. Highly cited publications can thus be analysed to identify those exerting notable influence within the journal's corpus [95].

The objective of this research is to deepen the understanding of bibliometric data through advanced visualization techniques, primarily using VOSviewer [94] and Bibliometrix [5]. VOSviewer processes bibliographic information extracted from databases such as WoS to generate graphical representations based on various bibliometric methods. The analysis in this study focuses on three main techniques: co-citation analysis [87], bibliographic coupling [51], and keyword co-occurrence analysis [15].

Co-citation analysis examines documents jointly cited by the same publications, revealing intellectual linkages between scholarly works [87]. This approach can be applied to different entities—such as documents, journals, or authors—emphasizing the relationships established through shared citations. In VOSviewer visualizations, the size of each node (circle) represents the number of citations, while the links between nodes reflect the strength of co-citation connections.

Bibliographic coupling, conversely, analyses documents that cite common sources [51]. This technique can be applied to multiple units of analysis, including documents, authors, institutions, and countries. In VOSviewer visualizations, node size represents the number of publications associated with each entity, whereas link thickness indicates the intensity of shared bibliographic references, thus mapping collaboration and thematic proximity within the dataset.

Keyword co-occurrence analysis identifies the simultaneous appearance of keywords within the same publications [30]. This study focuses on author-provided keywords, typically listed on publication title pages, to ensure thematic precision. Although abstract-based keyword extraction is an alternative, it often produces generic terms (e.g., method, model, analysis) that lack discriminative power. In VOSviewer, the size of each keyword node represents its frequency of occurrence, while the connecting lines denote co-occurrence relationships, collectively illustrating the thematic structure of the journal.

Several databases provide bibliographic and citation data. The most widely used general databases include WoS, Scopus, and Google Scholar, while specialized databases cover particular disciplines—such as PubMed and MathSciNet [25]. The bibliographic data for this study were retrieved from the WoS Core Collection [18].

Table 1 summarizes the stages and parameters of the bibliometric review. The study follows the systematic procedures and rationale established by the Scientific Procedures and Rationales for Systematic Literature Reviews (SPAR-4-SLR) protocol [76].

The data were retrieved from the WoS Core Collection for the period 2004–2024 using the following search query: Publication Title = “Knowledge and Information Systems”. The search was conducted on April 15, 2025, and all records were downloaded on the same date to ensure consistency. The dataset was then cleaned by removing duplicate records, excluding early access documents, and restricting the document types to articles and reviews. Author names and institutional affiliations were inspected for inconsistencies and standardized where necessary using the data cleaning functions available in Bibliometrix and manual verification.

An initial search for “Knowledge and Information Systems” in the “Publication Titles” field yielded 2,477 records. This dataset was refined through a series of filters to ensure analytical consistency. The first filter, Show Final Publication Year: Exclude 2025, reduced the total to 2,366 records. Applying the second filter, Document Types: Article and Review, narrowed the results to 2,340 records. A final filter, Exclude Early Access, produced 2,270 records.

To ensure comprehensive coverage of the journal’s publication history, 115 additional documents from the WoS Cited Reference Search and the KAIS official website were incorporated, covering the period 1999–2003, since KAIS was not indexed in WoS before 2004. Consequently, a total of 2,385 documents were included in the final dataset for analysis.

4 Results

This section presents a synthesis of the bibliometric findings for KAIS. It begins with an overview of the journal’s publication and citation patterns, followed by an examination of its most influential articles, key citing works, and the main contributors in terms of authors, institutions, and countries.

4.1 Publication and citation structure of KAIS

Figure 1 illustrates the annual evolution of the 2,385 articles and reviews published in KAIS between 1999 and 2024. Overall, the journal exhibits a clear long-term growth trajectory, reflecting its increasing recognition within the fields of computer science and information systems. Publication output grew gradually during the early years, stabilized in the mid-2000s, and accelerated markedly after 2015, culminating in the highest annual output in 2024.

In contrast, 2021 shows a temporary decline, with 117 publications—a reduction of approximately 27% compared to 2020. This drop may be partly associated with the global disruptions caused by the COVID-19 pandemic, which affected academic workflows, peer-review timelines, and research productivity across many disciplines. It may also reflect temporary editorial or operational adjustments during that period, although the available data do not allow a definitive causal interpretation.

Following this brief slowdown, KAIS experienced a strong rebound from 2022 onwards, with publication levels rising again after the 2021 decline. This recovery coincides with the growing prominence of research topics such as artificial intelligence, big data, and intelligent systems, which have increasingly shaped the journal’s thematic profile.

Table 1 Procedure of the study based on the SPAR-4-SLR protocol

Activity		
Assembling	Identification	Domain: <i>Knowledge and Information Systems</i> (KAIS) Research objective: Present a bibliometric overview of KAIS Research questions: (1) What are the publication and citation characteristics of KAIS? (2) Which are the most cited articles of KAIS according to WoS? (3) Which documents have been cited more frequently in KAIS? (4) Who are the most productive authors, institutions and countries of KAIS? (5) How is the co-citation and bibliographic coupling structure of KAIS? (6) How is the keyword and topical structure of KAIS? (7) How have the main research themes in KAIS evolved across different AI paradigms over time?
	Acquisition	Search platform: WoS Core Collection Search period: 2004–2024. The search was retrieved in April 2025 Source Title: Knowledge and Information Systems. Results = 2,477 Cited Reference Search and webpage of KAIS: 1999–2003. Results = 115 Number of citing articles: 41,103 documents Computer software: VOSviewer, <i>Bibliometrix</i>, <i>Biblioshiny</i> and Microsoft Excel
Arranging	Purification	Filter 1. Shows Final Publication Year and Exclude 2025: 2,366 documents Filter 2. Article & Review 2,340 documents Filter 3: Exclude Early Access: 2,270 documents Total number including documents from 1999–2003: 2,385 documents Number of citing articles after filters 1–3: 28,411 citing articles Additional results: JCR (WoS) and SciVal (Scopus)
Assessing	Evaluation	Performance analysis: (1) Rankings based on bibliometric indicators according to the number of articles and citations (2) Other bibliometric indicators: Cites per paper, the <i>h</i>-index, cites per year, papers and cites per capita, geographical analysis and temporal classification Graphical analysis: (1) Visualization with VOSviewer and Bibliometrix software (2) Techniques: Co-citation, bibliographic coupling, and co-occurrence of author keywords (3) Methods: Network analysis, thematic map and word cloud Agenda proposal method: (1) Present a bibliometric review of KAIS (2) Study the main results and findings of KAIS (3) Analyse questions for future research
	Reporting	Reporting conventions: Explanations with numbers, tables and figures Limitations: (1) The drawbacks of the WoS Core Collection database (2) The bibliometric information is from April 2025 but changes through time (3) Comparing subtopics with different publication and citation patterns

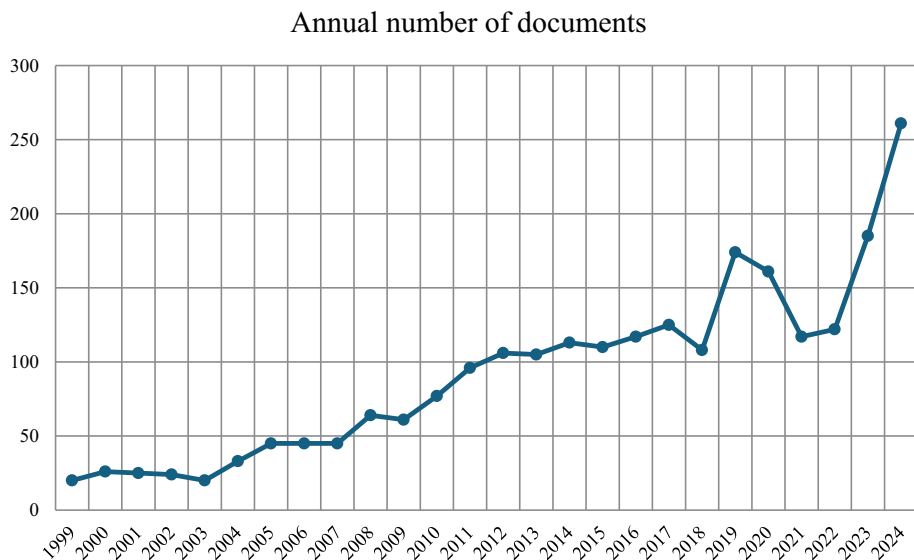


Fig. 1 Annual number of papers published in KAIS

Overall, the publication trajectory highlights KAIS's steady expansion and its continued relevance as a central venue for research in knowledge and information systems.

In addition to publication volume, citation performance provides a clearer indication of the journal's scholarly influence. Table 2 summarizes the citation distribution across all published documents. The results show that the vast majority of KAIS papers have been cited at least once (92.9%), and nearly two-thirds have received five or more citations, indicating strong integration into the global research community. Moreover, a non-negligible share of articles has achieved high citation impact, with over 8% exceeding 50 citations.

The year 2008 stands out as particularly influential. Among the 64 articles published that year, five have each received more than 100 citations, including the journal's most cited paper, which has accumulated over 3,000 citations. As a result, 2008 represents the single year with the highest aggregate citation impact in KAIS's history.

Overall, the citation distribution demonstrates that KAIS not only sustains a high publication output but also consistently produces papers with long-term scholarly influence, confirming its role as a key venue for both methodological and applied advances in knowledge and information systems.

Figure 2 presents the citation distribution of all KAIS publications using a box-whisker plot [91]. Each box represents the citation range for papers published in a specific year, highlighting the 25th, 50th (median), and 75 percentiles, together with the interquartile range (IQR), the average citations per paper, and the minimum and maximum values. Outliers exceeding 300 citations are depicted as yellow dots, based on citation data from the Web of Science Core Collection (December 2024) [30].

The overall shape of the distributions skews toward the upper end of the scale, indicating that a subset of articles consistently attracts disproportionately high citation counts—a characteristic pattern in established scientific journals [8]. These outliers, represented by yellow markers, correspond to high-impact papers that have exerted strong and lasting influence within the academic community.

Table 2 Annual citation structure of KAIS

Year	TP	TC	≥ 200	≥ 100	≥ 50	≥ 20	≥ 10	≥ 5	≥ 1	T50
1999	20	1,799	3	3	4	4	7	9	20	3
2000	26	350	0	0	0	6	16	18	26	0
2001	25	1,493	1	1	5	11	15	18	25	1
2002	24	350	0	0	2	5	8	14	24	0
2003	20	284	0	1	1	4	7	12	20	0
2004	33	1,061	2	3	5	13	20	25	32	1
2005	45	2,476	2	3	8	19	27	37	44	2
2006	45	1,359	0	2	8	21	33	40	45	1
2007	45	2,028	2	5	13	21	28	38	44	2
2008	64	5,542	3	5	10	27	48	60	64	2
2009	61	1,335	1	3	6	21	39	51	61	1
2010	77	2,292	1	4	11	34	56	69	74	4
2011	96	2,812	1	6	14	43	72	88	95	4
2012	106	3,242	2	4	11	42	63	87	102	2
2013	105	3,279	3	5	14	38	61	85	101	4
2014	113	3,243	1	4	9	30	55	85	107	2
2015	110	3,724	3	6	12	39	75	93	109	5
2016	117	2,353	0	2	11	34	62	92	116	1
2017	125	3,629	2	6	16	43	70	96	121	2
2018	108	2,518	0	4	11	42	64	85	104	1
2019	174	2,788	1	6	10	31	73	113	169	1
2020	161	2,070	0	3	7	22	54	100	152	2
2021	117	1,193	1	1	2	16	33	61	102	1
2022	122	1,363	0	2	2	17	32	65	108	2
2023	185	1,027	0	0	2	9	29	69	153	0
2024	261	709	0	1	1	2	6	33	198	0
Total	2,385	54,319	29	80	195	594	1,053	1,543	2,216	44
%	100		1.2	3.4	8.2	24.9	44.2	64.7	92.9	1.8

TP and TC = Total papers and citations; ≥ 200, ≥ 100, ≥ 50, ≥ 20, ≥ 10, ≥ 5, ≥ 1 = Number of papers with equal or more than 200, 100, 50, 20, 10, 5 and 1 citations; T50 = Number of documents in the Top 50 of Table 4

Notably, 2008 emerges as the most prominent year, showing the highest total citation count. This is explained by the presence of several widely referenced papers—five exceeding 100 citations each, including the journal's all-time most cited article. The concentration of citations in 2008 underscores the long-term influence of foundational works published during that period, which continue to anchor subsequent developments in data mining and information systems research.

To complement the citation analysis, Table 3 presents the evolution of KAIS performance metrics from 2014 to 2023, based on data from the Journal Citation Reports (JCR, [18]) and Scopus [85]. The results indicate that KAIS has maintained a stable and upward trajectory

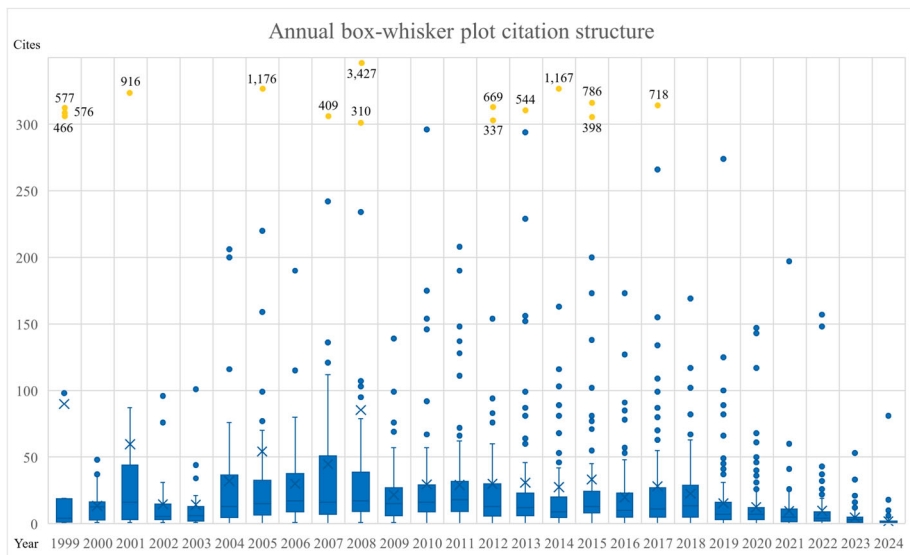


Fig. 2 Annual box-whisker plot citation structure of all papers published in KAIS

across multiple indicators, consistently ranking in the mid-to-high tier of journals in the Computer Science—Artificial Intelligence and Information Systems categories.

The journal's impact factor and percentile ranking demonstrate sustained visibility and influence, confirming KAIS's solid position within the scientific community. In 2018, the immediacy index reached its peak at 0.75, reflecting rapid citation uptake of newly published articles. Although a modest decline occurred in 2021, the subsequent rebound—with a 0.5-point increase in citation rate—suggests a recovery in scholarly attention and a return to steady performance. Moreover, the five-year impact factor peaked in 2019, indicating that KAIS maintains long-term citation relevance beyond short-term publication cycles.

Another relevant metric, the Article Influence Score (AIS) [10], captures the average relative impact of a journal's articles compared with the overall JCR database. During 2018–2020, KAIS recorded its highest AIS values, though these remained slightly below the global average ($\text{AIS} < 1.0$). This outcome reflects the journal's selective niche positioning—focusing on technically specialised, high-quality contributions that, while not always achieving mass citation volume, exert substantial influence within their subfields.

Overall, these bibliometric indicators reveal that KAIS has achieved consistent scholarly recognition, resilience in citation performance, and increasing alignment with the upper quartile of leading journals in its discipline.

Beyond these quantitative measures, the sustained growth of KAIS across both WoS and Scopus indexing systems reflects its increasing international prestige and citation visibility. The journal's consistent placement within the Q2–Q1 quartiles of computer science categories underscores its recognition among the leading venues for disseminating advances in knowledge systems, artificial intelligence, and information management. The rising trend in global citations—particularly from Asia–Pacific, North America, and Europe—demonstrates that KAIS has evolved into a truly international platform, fostering interdisciplinary dialogue across technical and applied domains. This trajectory not only consolidates its reputation as a

Table 3 Analysis of KAIS in the JCR of the WoS and Scopus

Year	TCJ	IF	5YIF	ImI	AIS	RAI	PAI	RIS	PIS	CS	PB
2014	1,565	1.78	2.01	0.20	0.62	46/123	63.01	30/139	78.78	5.1	88-HA
2015	1,559	1.70	1.72	0.29	0.65	54/130	58.85	41/144	71.88	4.5	81-HA
2016	2,146	2.00	2.02	0.28	0.60	57/133	57.52	67/146	54.45	4.6	80-HA
2017	2,343	2.24	1.99	0.36	0.66	49/132	63.26	56/148	62.5	4.4	80-HA
2018	3,006	2.39	2.54	0.75	0.72	64/134	52.61	70/155	55.16	5.4	82-HA
2019	3,242	2.93	2.68	0.43	0.75	52/137	62.41	57/156	63.78	5.2	78-HA
2020	4,008	2.82	3.16	0.54	0.77	65/139	53.6	79/161	51.24	5.8	79-IS
2021	4,789	2.53	3.20	0.23	0.65	90/145	38.28	110/164	33.23	5.9	75-IS
2022	5,010	2.7	3.3	0.5	0.62	95/145	34.8	105/158	33.9	6.1	75-IS
2023	4,710	2.5	2.6	0.5	0.57	106/197	46.4	119/250	52.6	5.7	66-IS
2024	5,235	3.1	2.8	0.5	0.591	95/204	53.7	111/258	57.2	5.7	70-IS

TCJ = Total citations in the JCR; IF = Impact factor; 5YIF = 5-year impact factor; ImI = Immediacy index; AIS = Article Influence Score; RAI = Ranking in the WoS category of Computer Science, Interdisciplinary Applications; PAI = Journal impact factor percentile in Computer Science, Interdisciplinary Applications; RIS = Ranking in the WoS category of Engineering, Industrial; PIS = Journal impact factor percentile in Engineering, Industrial; CS = CiteScore; PB = Best Percentile in Scopus among the four categories that indexes KAIS (AI = Artificial Intelligence; HA = Hardware and Architecture Software; HC = Human-Computer Interaction; IS = Information Systems; S = Software)

high-impact journal but also reinforces its role in shaping the intellectual and methodological direction of the knowledge and information systems community.

4.2 Influential papers in KAIS

Table 4 presents the 50 most influential papers published in KAIS between 1999 and 2024. Among them, Wu et al. [98] remains the journal's most impactful contribution. By synthesizing the "Top 10 algorithms in data mining", it codified the intellectual DNA of the field. The paper's enduring influence derives from its pedagogical clarity and cross-domain reach, serving both as an educational resource and as a conceptual anchor. It introduced a shared vocabulary—C4.5, SVMs, PageRank, EM, AdaBoost, among others—that continues to appear in textbooks, benchmark frameworks, and data mining curricula. Its main limitation is temporal: published before the deep learning revolution, it omits more recent paradigms such as reinforcement learning, representation learning, and self-supervised models, highlighting the need for a contemporary update reflecting the post-2015 landscape. Notably, this paper is currently ranked 5th among the most-cited articles of 2008 in the WoS category "Computer Science, Information Systems" (5th of 43,880 publications).

The second most-cited article, Keogh and Ratanamahatana [49], with 1,176 citations, introduced a rigorous framework for exact indexing of dynamic time warping (DTW), addressing long-standing challenges in time-series similarity search. By formalizing the efficiency–accuracy trade-off, the study established DTW as a foundational technique in pattern recognition, sensor data analytics, and anomaly detection. Its limitations—particularly regarding scalability and adaptation to multivariate, non-stationary series—have motivated later developments in adaptive indexing and streaming time-series models, ensuring its continuing relevance in temporal data mining. Together with Keogh et al. (2001) [50] and [49], this line of research reshaped time-series data mining. The 2001 work on dimensionality reduction enabled scalable similarity search, while the 2005 paper—provocatively titled "Clustering of time-series subsequence is meaningless"—exposed methodological flaws in prior studies [48]. The latter's influence is both corrective and foundational, promoting experimental rigor and statistical validity that have become hallmarks of KAIS publications.

Yang and Leskovec (2015) [101] and Aminikhanghahi and Cook (2017) [4] exemplify KAIS's influence on network science and temporal analysis. Yang and Leskovec (2015) [101] introduced ground-truth community definitions, establishing robust evaluation metrics that remain standard in graph learning. Aminikhanghahi and Cook (2017) [4] provided a comprehensive survey of change-point detection, synthesizing decades of algorithms into a coherent taxonomy widely cited in domains such as bioinformatics, IoT analytics, and financial modeling. Together, these works highlight KAIS's role as a repository of methodological consolidation and cross-disciplinary knowledge integration.

Kamiran and Calders (2012) [47] pioneered data preprocessing for fairness, marking one of the earliest algorithmic fairness frameworks in data mining. Its influence extends beyond computer science, shaping research in social computing, AI ethics, and policy studies. The approach's simplicity—based on reweighting and relabelling—facilitated broad adoption, though its dependence on group-level parity metrics limits its applicability in contexts requiring causal or counterfactual fairness. Coello (1999)[22], a seminal survey of multi-objective evolutionary optimization, also ranks among KAIS's most cited works. It unified decades of algorithmic research into a comprehensive taxonomy that has since informed hybrid evolutionary–machine learning systems. Despite its age, it continues to be cited for its methodological clarity and historical perspective.

Table 4 The 50 most cited documents in KAIS

R	TC	Title	Author/s	Year	C/Y
1	3,427	Top 10 algorithms in data mining	Wu, XD; Kumar, V; Quinlan, JR; Ghosh, J; Yang, Q; Motoda, H; McLachlan, GJ; Ng, A; Liu, B; Yu, PS; Zhou, ZH; Steinbach, M; Hand, DJ; Steinberg, D	2008	190.39
2	1,176	Exact indexing of dynamic time warping	Keogh, E; Ratanamahatana, CA	2005	56
3	1,167	Explaining prediction models and individual predictions with feature contributions	Strumbelj, E; Kononenko, I	2014	97.25
4	916	Dimensionality reduction for fast similarity search in large time series databases	Keogh, E; Chakrabarti, K; Pazzani, M; Mehrotra, S	2001	38.17
5	786	Defining and evaluating network communities based on ground-truth	Yang, J; Leskovec, J	2015	71.45
6	718	A survey of methods for time series change point detection	Aminikhanghahi, S; Cook, DJ	2017	79.78
7	669	Data preprocessing techniques for classification without discrimination	Kamiran, F; Calders, T	2012	47.79
8	577	A comprehensive survey of evolutionary-based multiobjective optimization techniques	Coello, CAC	1999	22.19
9	576	Data preparation for mining World Wide Web browsing patterns	Cooley, R; Mobasher, R; Srivastava, J	1999	22.15
10	544	A review of feature selection methods on synthetic data	Bolón-Canedo, V; Sánchez-Marono, N; Alonso-Betanzos, A	2013	41.85
11	466	Agent-based systems for intelligent manufacturing: a state-of-the-art survey	Shen, W; Norrie, DH	1999	17.92
12	409	Stability of feature selection algorithms: a study on high-dimensional spaces	Kalousis, A; Prados, J; Hilario, M	2007	21.53
13	398	Self-labeled techniques for semi-supervised learning: taxonomy, software and empirical study	Triguero, I; García, S; Herrera, F	2015	36.18

Table 4 (continued)

R	TC	Title	Author/s	Year	C/Y
14	337	SMOTE-RSB*: a hybrid preprocessing approach based on oversampling and undersampling for high imbalanced data-sets using SMOTE and rough sets theory	Ramentol, E; Caballero, Y; Bello, R; Herrera, F	2012	24.07
15	310	Comparison of descriptor spaces for chemical compound retrieval and classification	Wale, N; Watson, IA; Karypis, G	2008	17.22
16	296	Semi-supervised learning by disagreement	Zhou, ZH; Li, M	2010	18.5
17	294	Transfer learning for activity recognition: a survey	Cook, D; Feuz, KD; Krishnan, NC	2013	22.62
18	274	A survey of sentiment analysis in social media	Yue, L; Chen, WT; Li, X; Zuo, WL; Yin, MH	2019	39.14
19	266	Recent advances in feature selection and its applications	Li, Y; Li, T; Liu, H	2017	29.56
20	242	Supervised tensor learning	Tao, DC; Li, XL; Wu, XD; Hu, WM; Maybank, SJ	2007	12.74
21	234	Random walk with restart: fast solutions and applications	Tong, HH; Faloutsos, C; Pan, JY	2008	13
22	229	A survey on instance selection for active learning	Fu, YF; Zhu, XQ; Li, B	2013	17.62
23	220	Clustering of time-series subsequences is meaningless: implications for previous and future research	Keogh, E; Lin, J	2005	10.48
24	208	Application of correlation coefficient to interval-valued intuitionistic fuzzy multiple attribute decision-making with incomplete weight information	Wei, GW; Wang, HJ; Lin, R	2011	13.87
25	206	Ontology evolution: Not the same as schema evolution	Noy, NF; Klein, M	2004	9.36
26	200	A survey on data stream clustering and classification	Nguyen, HL; Woon, YK; Ng, WK	2015	18.18
27	200	Ontologies: How can they be built?	Pinto, HS; Martins, JP	2004	9.09
28	197	Model complexity of deep learning: a survey	Hu, X; Chu, LY; Pei, J; Liu, WQ; Bian, J	2021	39.4
29	190	An overview on subgroup discovery: foundations and applications	Herrera, F; Carmona, CJ; González, P; del Jesus, MJ	2011	12.67

Table 4 (continued)

R	TC	Title	Author/s	Year	C/Y
30	190	Using discriminant analysis for multi-class classification: an experimental investigation	Li, T; Zhu, SH; Ogihara, M	2006	9.5
31	175	Boosting support vector machines for imbalanced data sets	Wang, BX; Japkowicz, N	2010	10.94
32	173	A survey on indexing techniques for big data: taxonomy and performance evaluation	Gani, A; Siddiqua, A; Shamshirband, S; Hanum, F	2016	17.3
33	173	Evaluating link prediction methods	Yang, Y; Lichtenwalter, RN; Chawla, NV	2015	15.73
34	169	Core techniques of question answering systems over knowledge bases: a survey	Diefenbach, D; Lopez, V; Singh, K; Maret, P	2018	21.13
35	163	A weighted voting framework for classifiers ensembles	Kuncheva, LI; Rodríguez, JJ	2014	13.58
36	159	Generative model-based document clustering: a comparative study	Zhong, S; Ghosh, J	2005	7.57
37	157	Interpretable deep learning: interpretation, interpretability, trustworthiness, and beyond	Li, XH; Xiong, HY; Li, XJ; Wu, XY; Zhang, X; Liu, J; Bian, J; Dou, DJ	2022	39.25
38	156	In-network outlier detection in wireless sensor networks	Branch, JW; Giannella, C; Szymanski, B; Wolff, R; Kargupta, H	2013	12
39	155	EFIM: a fast and memory efficient algorithm for high-utility itemset mining	Zida, S; Fournier-Viger, P; Lin, JCW; Wu, CW; Tseng, VS	2017	17.22
40	154	On the choice of the best imputation methods for missing values considering three groups of classification methods	Luengo, J; García, S; Herrera, F	2012	11
41	154	Extension of TOPSIS method for 2-tuple linguistic multiple attribute group decision making with incomplete weight information	Wei, GW	2010	9.63
42	152	Topic-aware social influence propagation models	Barbieri, N; Bonchi, F; Manco, G	2013	11.69
43	148	From distributed machine learning to federated learning: a survey	Liu, J; Huang, JZ; Zhou, Y; Li, XH; Ji, SL; Xiong, HY; Dou, DJ	2022	37

Table 4 (continued)

R	TC	Title	Author/s	Year	C/Y
44	148	The ClusTree: indexing micro-clusters for anytime stream mining	Kranen, P; Assent, I; Baldauf, C; Seidl, T	2011	9.87
45	147	Pythagorean fuzzy MULTIMOORA method based on distance measure and score function: its application in multicriteria decision making process	Huang, C; Lin, MW; Xu, ZH	2020	24.5
46	146	Tracking recurring contexts using ensemble classifiers: an application to email filtering	Katakis, I; Tsoumakas, G; Vlahavas, I	2010	9.13
47	143	A survey on influence maximization in a social network	Banerjee, S; Jenamani, M; Pratihari, DK	2020	23.83
48	139	Characterization and evaluation of similarity measures for pairs of clusterings	Pfützner, D; Leibbrandt, R; Powers, D	2009	8.18
49	138	Class imbalance revisited: a new experimental setup to assess the performance of treatment methods	Prati, RC; Batista, GEAPA; Silva, DF	2015	12.55
50	137	The k-anonymity and l-diversity approaches for privacy preservation in social networks against neighborhood attacks	Zhou, B; Pei, J	2011	9.13

R = Rank; TC = Total citations; C/Y = Cites per year

Cooley et al. (1999) [23] established one of the first web usage mining frameworks, integrating data cleaning, sessionization, and pattern discovery—processes that now underpin modern big data pipelines and recommendation systems. Its longevity stems from its structured methodology and practical orientation, which continue to inform web analytics and user behaviour modeling. Bolón-Canedo et al. (2013) [13] completes the top ten, offering a comprehensive evaluation of feature selection methods using synthetic data. Its main contribution lies in addressing the reproducibility crisis in feature selection through controlled benchmarks, a contribution that remains crucial for studies on high-dimensional data analysis. Nevertheless, the rise of deep feature learning has shifted current attention toward end-to-end representation learning, prompting renewed interest in integrating feature selection with deep architectures.

Regarding citations per year, Wu et al. [106] again leads, with an average of 190.39 citations annually, followed by Strumbelj and Kononenko (2014) [88], averaging 97.25 citations per year. The latter advanced model interpretability by introducing feature contribution explanations, a model-agnostic framework that directly inspired LIME and SHAP. Its strength lies in its transparency and generalizability across learning models, and its influence has grown as interpretability became a core component of trustworthy AI. However, full integration with

deep neural architectures and causal reasoning frameworks remains an open research direction. This paper currently ranks as the 18th most-cited article of 2014 in the WoS category “Computer Science, Information Systems” (18th of 53,067 publications).

The Table 4 presents the 50 most cited papers published in KAIS. The analysis of these papers reveals several defining patterns in the journal’s intellectual development.

First, the most influential contributions are predominantly methodological. Many of the top-ranked papers either introduce foundational algorithms or provide comprehensive surveys that consolidate knowledge across subfields. Among the Top 50 papers, 14 correspond to survey or review articles, while the remaining 36 present algorithmic or methodological contributions. Although algorithmic papers are numerically dominant, several surveys appear among the highest-ranked and fastest-growing papers in terms of citations per year, suggesting that knowledge synthesis plays a crucial role in long-term citation impact within KAIS.

Second, temporal and structural data analysis emerges as a central intellectual thread. Highly cited papers on time-series analysis, graph mining, and stream processing consistently appear among the top ranks. This indicates that KAIS has maintained a strong methodological identity in dynamic and relational data analysis, distinguishing it from more application-driven journals.

Third, several influential papers address fairness, interpretability, and trustworthiness. These topics act as conceptual bridges between classical data mining and contemporary artificial intelligence, reflecting the journal’s adaptation to evolving scientific and societal challenges [89].

Fourth, a number of highly cited works focus on evaluation frameworks, benchmarking, and reproducibility. These contributions have established methodological standards that continue to guide current research in deep learning and data-centric AI.

Additional evidence of high-impact topics is provided by the KAIS Best Paper Awards, established in 2018 to recognize papers with strong citation performance three years after publication. Among the award-winning papers, five also appear in the Top 50 most cited articles identified in Table 4, including works on network community detection [101], time-series change-point detection [4], deep learning complexity [42], interpretable deep learning [63], and fuzzy multi-criteria decision making [43].

Overall, the corpus of top-cited papers highlights KAIS’s enduring role in methodological consolidation, cross-domain applicability, and responsible knowledge discovery.

Investigating the most cited works in articles published in KAIS also aids in analysing the journal’s contributions and influence. Table 5 lists the Top 30 most cited documents from KAIS publications in the WoS Core Collection database. It provides a summary of the articles referenced in this journal, which assists in grasping the journal’s topical impact and influence.

The analysis of the most frequently cited external references reveals the methodological foundations upon which KAIS research is built. The most cited works correspond to three major pillars of modern data science: algorithmic pattern discovery, statistical evaluation and experimental methodology, and probabilistic and representation-based modeling.

Agrawal et al. [2] introduced the Apriori algorithm and formalised the problem of association rule mining, establishing a scalable framework for uncovering relational patterns in large transactional databases. Within KAIS, this work is cited not merely as a historical milestone but as a conceptual foundation for subsequent studies in pattern discovery, frequent itemset mining, recommender systems, and knowledge extraction from temporal or uncertain data. Its enduring relevance underscores how classical algorithmic paradigms continue to evolve alongside new computational contexts such as streaming and distributed environments.

Demšar [24] addressed a critical methodological gap by proposing a rigorous statistical framework for comparing classifiers across datasets. Through a systematic assessment

Table 5 Top 30 most cited documents in KAIS publications

Rank	First author	Reference	Year	Vol	Page	Type	TC
1	Agrawal R	SIGMOD Record	1993	v22	p207	A	86
2	Demsar J	J Mach Learn Res	2006	v7	p1	A	85
3	Blei DM	J Mach Learn Res	2003	v3	p993	A	79
4	Wu XD	Knowl Inf Syst	2008	v14	p1	A	63
5	Agrawal R	P 1994 Int C Ver Lar	1994		p487	C	62
6	Vaswani A	Adv Neur In	2017	v30		C	60
7	Mikolov T	P Int C Learning Rep	2013			C	49
8	Agrawal R	Proc Int Conf Data	1995	p3		C	47
9	Dempster AP	J Roy Stat Soc B Met	1977	v39	p1	A	46
10	Devlin J	NAACL HLT 2019	2019	v1	p4171	C	45
11	King DB	Acs Sym Ser	2015	v1214	p1	C	45
12	Pennington J	Proc EMNLP	2014		p1532	C	43
13	Zadeh LA	Inform Control	1965	v8	p338	A	40
14	Breiman L	Machine Learning	2001	v45	p5	A	39
15	Grover A	KDD' 16: Proc 22nd ACM SIGKDD	2016		p855	C	38
16	Quinlan JR	C4. 5: Programs Machine Learning	2014			B	38
17	Graves A	Stud Comput Intell	2012	v385	p1	B	37
18	Jain AK	ACM Comput Surv	1999	v31	p264	A	37
19	Miller GA	Commun ACM	1995	v38	p39	A	36
20	Chawla NV	J Artif Intell Res	2002	v16	p321	A	35
21	Hamilton WL	Adv Neur In	2017	v30		C	34
22	Pan SJ	IEEE T Knowl Data En	2010	v22	p1345	A	34
23	Koren Y	Computer	2009	v42	p30	A	34
24	Deerwester S	J Am Soc Inform Sci	1990	v41	p391	A	34
25	Asuncion A	UCI Machine Learning Repository	2007			W	33
26	Cortes C	Mach Learn	1995	v20	p273	A	33
27	Fortunato S	Phys Rep	2010	v486	p75	A	31
28	He KM	Proc CVPR IEEE	2016		p770	C	30
29	Kipf TN	ICLR	2016			C	30
30	Perozzi B	KDD' 14: Proc 20th ACM SIGKDD	2014		p701	C	30

TC = Total citations; A = Article; B = Book; C = Conference proceedings; W = Webpage

of parametric (t-test, ANOVA) and non-parametric (Wilcoxon, Friedman) tests, the study demonstrated that non-parametric methods provide more reliable results when standard distributional assumptions are violated. In KAIS, this reference is most often cited in the methodology sections of empirical papers, validating comparative analyses and reinforcing the journal's commitment to reproducibility and statistical integrity.

Blei et al. [12] extended the methodological horizon of KAIS by introducing Latent Dirichlet Allocation (LDA), a probabilistic generative model that revolutionised topic modelling and textual knowledge discovery. Within the journal's corpus, this work bridges information retrieval, semantic analysis, and knowledge representation, serving as a foundational citation in studies on ontology evolution, text mining, and knowledge graph construction. Its Bayesian framework exemplifies the journal's emphasis on interpretable, model-based learning—a thematic link between early data mining traditions and the probabilistic reasoning paradigms that underpin contemporary AI research.

Classic contributions in association rule mining, classifier comparison, and topic modeling dominate the citation landscape. These works provide general-purpose methods and evaluation frameworks that support a wide range of research topics within KAIS.

This pattern indicates that KAIS publications rely heavily on methodological cornerstones rather than narrow domain-specific references. The prominence of such foundational works confirms the journal's orientation toward generalizable algorithms, evaluation standards, and knowledge modeling techniques.

4.3 Citing articles of KAIS

Another factor to consider is to investigate the origin of the citations that KAIS receives. For this analysis, we focus on articles that have cited the journal, meaning those that mention it at least once in their reference list. The search includes only records from 2004 onward, which is when the journal was indexed in WoS. Initially, the study selects “Citing articles” in the Citation Reports of WoS for the 2,270 documents of KAIS, obtaining 41,103 documents (i.e., citing articles). Next, excluding 2025 narrows it down further to 39,629 documents. A second filter, “Document Type: Article and Review”, brings the countdown to 29,048 documents, and finally, by excluding “Early Access” and “Retracted” articles, the total number of citing articles becomes 28,411.

Figure 3 illustrates the annual trend in published articles citing KAIS. The data show a significant upward trend starting in 2017, with 1,452 cited articles, which increased to 4,006 by 2024. This sustained growth reflects three interrelated factors. First, the broadening thematic scope of KAIS—from classical knowledge systems to contemporary fields such as data science, deep learning, and AI interpretability—has expanded its relevance to a wider scientific audience. Second, the journal's interdisciplinary positioning at the intersection of computer science, engineering, decision systems, and management has made it a preferred outlet for research seeking both technical depth and applied impact. Third, the increase parallels the global expansion of AI and data-driven research, where bibliometric analyses, knowledge discovery, and decision-support methodologies—core areas in KAIS—have become central to scientific innovation.

Consequently, the upward trend in citing articles does not merely reflect quantitative growth but indicates KAIS's evolving role as a methodological and conceptual reference point across disciplines such as computer science, telecommunications, mathematics, operations research, automation, environmental sciences, and business analytics. The journal's increasing citation visibility thus underscores its integration into the mainstream discourse of intelligent and knowledge-based systems research.

Table 6 outlines the top universities and countries that frequently cite the journal, based on the total number of publications documented in the WoS Core Collection database. Notably, the Chinese Academy of Sciences holds the top position for the highest number of citations to KAIS articles, with institutions from France, China, and Australia, following closely among

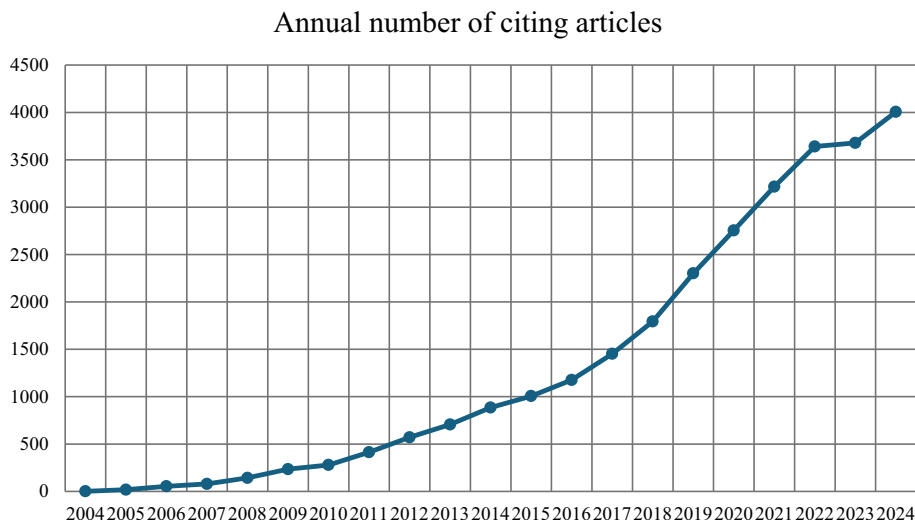


Fig. 3 Annual number of citing articles of KAIS

the top five. At the national level, China, the United States, India, the United Kingdom, and Australia are the countries that most often publish research citing KAIS.

These patterns reveal both the geographical diffusion and the epistemic centrality of KAIS within the global research landscape. The dominance of Chinese institutions mirrors the country's substantial investment in artificial intelligence, data analytics, and knowledge engineering, fields that align closely with KAIS's thematic focus. Similarly, the strong presence of the United States and the United Kingdom reflects the journal's continued influence in computational intelligence and information systems, while emerging citation activity from India and Australia indicates the global democratization of research capacity in data science and intelligent systems.

Overall, the concentration of citing institutions in technologically advanced economies highlights KAIS's international recognition as a methodological and applied reference point in knowledge discovery, machine learning, and decision-support systems. The journal's growing citation footprint across regions underscores its integration into mainstream AI and information systems research, affirming its global relevance and multidisciplinary reach.

Table 7 presents the journals and subject fields that most frequently cite KAIS. The leading citing sources are large, application-oriented journals, including *IEEE Access* and *Expert Systems with Applications*, together with KAIS itself. This pattern indicates that the journal's contributions are widely used in applied artificial intelligence, intelligent decision support, and knowledge-based systems research. In terms of subject areas, most citing articles fall within Computer Science, Engineering, Telecommunications, Mathematics, and Operations Research & Management Science, confirming the journal's multidisciplinary reach. The notable level of self-citations also suggests a cohesive research community, where newer studies frequently build upon earlier KAIS contributions.

The diversity of subject categories citing KAIS underscores its position at the intersection of foundational and applied research. Citations from fields such as telecommunications and operations research suggest that KAIS methodologies are being adapted beyond traditional computer science contexts, particularly in network optimization, cyber-physical systems,

Table 6 Citing articles of KAIS: Universities and countries

R	University	TP	Country	TP
1	Chinese Academy of Sciences	732	PR China	10,473
2	CNRS France	372	USA	4,350
3	Harbin Institute of Technology	346	India	2,135
4	Tsinghua U	274	UK	1,528
5	U Technology Sydney	271	Australia	1,509
6	Zhejiang U	270	Spain	1,451
7	U Electronic Science Tech China	254	Canada	1,043
8	Hefei U Technology	245	Italy	1,031
9	Islamic Azad U	239	Germany	1,028
10	U Granada	233	Iran	1,009
11	U Chinese Academy Sciences	228	South Korea	1,006
12	Wuhan U	225	France	907
13	Xidian U	212	Saudi Arabia	856
14	King Abdulaziz U	201	Taiwan	701
15	National U Defense Tech China	196	Brazil	622
16	China U Geosciences	195	Pakistan	616
17	Nanyang Technological U	195	Turkey	587
18	Beihang U	192	Japan	565
19	Shanghai Jiao Tong U	192	Malaysia	521
20	Dalian U Technology	185	Singapore	434
21	Sichuan U	171	Poland	425
22	Hong Kong Polytechnic U	169	Netherlands	414
23	Tongji U	169	Greece	382
24	Shenzhen U	166	Vietnam	325
25	Nanjing U	164	Egypt	324
26	Southeast U China	163	Belgium	319
27	Central South U	160	Portugal	285
28	Huazhong U Science Technology	160	Norway	281
29	Jilin U	160	Finland	252
30	Xi An Jiaotong U	158	Mexico	246

R = Rank; TP = Total papers

and data-driven management science. This disciplinary breadth confirms KAIS's growing reputation as a cross-domain reference source that bridges data-centric computation, decision analytics, and knowledge representation in both academic and industrial research.

4.4 Leading authors, institutions and countries

Another important dimension of the bibliometric analysis concerns identifying the most productive authors, institutions, and countries contributing to KAIS. This section first examines the authors with the highest number of publications, as summarized in Table 8, which ranks

Table 7 Citing articles of KAIS: Journals and subject areas

R	Journal	TP	Research area	TP
1	IEEE Access	1,275	Computer Science	18,492
2	Knowledge and Information Systems	1,077	Engineering	8,429
3	Expert Systems with Applications	817	Telecommunications	2,271
4	Knowledge Based Systems	602	Mathematics	1,520
5	Information Sciences	570	Operations Research Management Science	1,356
6	IEEE Trans Knowledge Data Engin	431	Science Technology Other Topics	1,328
7	Neurocomputing	409	Physics	1,131
8	Applied Sciences Basel	336	Chemistry	898
9	Applied Intelligence	335	Automation Control Systems	850
10	Applied Soft Computing	319	Environmental Sciences Ecology	799
11	Sensors	290	Business Economics	667
12	J Intelligent Fuzzy Systems	262	Materials Science	642
13	Scientific Reports	236	Information Science Library Science	566
14	Pattern Recognition	223	Instruments Instrumentation	561
15	Data Mining and Knowledge Discovery	221	Mathematical Computational Biology	549
16	Engin Applic Artificial Intelligence	220	Medical Informatics	443
17	PLOS One	213	Remote Sensing	415
18	Mathematics	180	Biochemistry Molecular Biology	379
19	Multimedia Tools and Applications	173	Geology	363
20	Neural Computing Applications	173	Energy Fuels	350
21	Electronics	169	Neurosciences Neurology	341
22	Soft Computing	169	Imaging Science Photographic Technology	338
23	ACM Trans Knowledge Discov Data	166	Transportation	265
24	Proceedings of the VLDB Endowment	163	Physical Geography	233
25	IEEE Trans Neural Netw Learning Syst	161	Biotechnology Applied Microbiology	208
26	ACM Computing Surveys	146	Water Resources	203
27	Intelligent Data Analysis	138	Construction Building Technology	192
28	Concurrency Comput Pract Experience	135	Health Care Sciences Services	163

Table 7 (continued)

R	Journal	TP	Research area	TP
29	J Supercomputing	135	Social Sciences Other Topics	156
30	Information Processing Management	128	Psychology	142

R = Rank; TP = Total papers

them by total output and, in cases of ties, by total citation count. Additional indicators, such as citations per paper and h-index values, are also reported to provide a more comprehensive understanding of author performance within the journal.

Philip S. Yu emerges as the most prolific author in KAIS, followed by Jian Pei and Jianzhong Li, reflecting the strong presence of leading researchers in data mining, large-scale data management, and intelligent information systems. In terms of citation impact, Philip S. Yu also ranks first, followed closely by Eamonn Keogh and Francisco Herrera, indicating a notable balance between productivity and scholarly influence among the journal's top contributors.

The geographical distribution of productive authors reveals a concentration in a few major research hubs, particularly China and the United States, with additional representation from Spain and Australia. Overall, the Top 30 authors span four continents, underscoring the journal's international reach and its role as a platform for cross-continental collaboration.

At the institutional level (Table 9), leading positions are occupied by research-intensive organizations with strong traditions in data mining and artificial intelligence. The Chinese Academy of Sciences ranks first in publication output, while IBM records the highest total citation count, highlighting the journal's integration of both academic and industrial research. Several highly cited universities also appear among the top institutions, reflecting sustained contributions from globally recognized research centers.

Overall, the results indicate that KAIS's most active contributors are concentrated in major research institutions with strong computational research capacity and international collaboration networks. This pattern reflects the journal's dual orientation toward methodological rigor and applied innovation.

To provide a broader perspective on the journal's global reach, the analysis also examines publication output at the country level, based on authors' institutional affiliations at the time of publication. Table 10 presents the 30 most productive and influential contributing countries.

China and the United States clearly dominate the ranking in terms of publication volume, followed by a second tier of active contributors including India, Australia, and Spain. In contrast, smaller but research-intensive countries such as Singapore exhibit high publication and citation efficiency relative to population size. Citation threshold analysis further shows that the majority of highly cited KAIS papers originate from China and the United States, while a wide range of European and Asian countries also appear among the top contributors.

Overall, these patterns highlight the progressive geographic diversification of KAIS, confirming its evolution into a globally distributed and collaborative research platform.

The dominance of Chinese and American contributions reflects broader trends in AI, data science, and intelligent systems, where both nations have made substantial investments in knowledge-driven innovation. At the same time, the strong presence of European and Asia-Pacific countries highlights the journal's role in facilitating cross-regional knowledge exchange.

Table 8 Most productive authors in KAIS

R	Author Name	University	Country	TP	TC	H	C/P	D1	D2	D3
1	Yu, Philip S	U Illinois Chicago	USA	29	4,132	15	142.48	9	15	5
2	Pei, Jian	Duke U	USA	18	624	9	34.67	3	12	3
3	Li, Jianzhong	Chinese Acad Sci	CHN	14	95	5	6.79	0	12	2
4	Keogh, Eamonn	U California Riverside	USA	13	1,841	9	141.62	5	6	2
5	Venkatesh, Svetha	Deakin U	AUS	12	158	8	13.17	0	12	0
6	Phung, Dinh Q	Deakin U	AUS	12	135	7	11.25	0	12	0
7	Gao, Hong	Zhejiang Normal U	CHN	12	80	5	6.67	0	10	2
8	Faloutsos, Christos	Carnegie Mellon U	USA	11	562	8	51.10	1	8	2
9	Herrera, Francisco	U Granada	SPA	10	1,397	10	139.7	0	9	1
10	Xiong, Hui	Rutgers U System	USA	10	206	9	20.6	2	7	2
11	Fournier-Viger, Philippe	Harbin Inst Tech	CHN	9	321	6	35.67	0	5	4
12	Kamel, Mohamed S	U Waterloo	CAN	9	217	8	25	2	7	0
13	Aggarwal, Charu C	IBM T J Watson Res Ctr	USA	9	159	7	17.67	1	7	1
14	Zou, Fengcai	Yunnan Agricultural U	CHN	8	370	7	46.25	3	5	0
15	Gionis, Aristides	Aalto U	FIN	8	105	6	13.13	2	5	1
16	Wang, Hongzhi	Harbin Inst Tech	CHN	8	57	4	7.13	0	5	3
17	Han, Xixian	Harbin Inst Tech	CHN	8	41	4	5.13	0	8	0
18	Li, Tao	Xiamen U	CHN	7	360	5	51.43	1	6	0
19	Li, Chaoqun	China U Geo-sciences	CHN	7	194	6	27.71	1	3	3
20	Jiang, Liangxiao	China U Geo-sciences	CHN	7	194	6	27.71	1	3	3
21	Webb, Geoff	Monash U	AUS	7	182	5	26	0	3	4

Table 8 (continued)

R	Author Name	University	Country	TP	TC	H	C/P	D1	D2	D3
22	Bailey, James	U Notting- ham	UK	7	169	6	24.14	2	3	2
23	Ventura, Sebastian	U Cordoba	SPA	7	127	7	18.14	0	5	2
24	Wang, Senzhang	Central South U	CHN	7	88	5	12.57	0	2	5
25	Han, Meng	North Minzu U	CHN	7	34	4	4.86	0	0	7
26	Tseng, Vincent S	Nat YM Chiao Tung U	TWN	6	330	5	55	0	6	0
27	Bonchi, Francesco	CENTAI Inst	ITA	6	228	4	38	1	5	0
28	Bajo, Javier	U Polytech Madrid	SPA	6	79	4	13.17	0	4	2
29	De Paz, Juan F	U Salamanca	SPA	6	67	4	11.17	0	4	2
30	Suzuki, Einoshin	Kyushu U	JAP	6	31	4	5.17	0	5	1

R = Rank; TP = Total papers; TC = Total citations; H = *h*-index; C/P = Cites per paper; D1 = TP between 1999–2009; D2 = TP 2010–2019; D3 = TP 2020–2024

From a strategic perspective, these patterns suggest that future growth in KAIS is likely to continue along international and interdisciplinary lines, driven by collaborations that integrate computational intelligence, information systems, and data-driven decision support. Thus, the country-level analysis provides valuable insights into how the journal's knowledge base has expanded globally and where emerging centres of scientific influence are forming.

To monitor the evolution of international participation in the journal, Table 11 presents the annual number of papers published in KAIS by the top 25 contributing countries over the past 21 years. This longitudinal perspective highlights the shifting landscape of research productivity and the global diffusion of knowledge within the KAIS community.

Historically, the United States dominated the journal's early publication record, accounting for the majority of contributions during its first decade. However, beginning in 2018, China surpassed the United States and has maintained its leadership since then, marking a major inflection point in the journal's geographical authorship distribution.

Between 2019 and 2024, China's output grew substantially—publishing 22 more papers than the United States in 2019, 25 more in 2021, and 71 more in 2024. This rapid acceleration aligns with China's national investment in artificial intelligence, big data, and information systems research, areas that directly align with KAIS's thematic scope. The trend suggests that China is likely to consolidate an even more dominant role in shaping future contributions to the journal. At the same time, India has shown consistent growth, emerging as a significant source of high-quality submissions in areas such as machine learning, decision support, and computational intelligence.

More broadly, most countries listed in Table 11 exhibit steady increases in annual publication output, a pattern that mirrors KAIS's growing international visibility and rising

Table 9 The 30 most productive and influential institutions in KAIS

R	Institution	Country	TP	TC	H	C/P	D1	D2	D3
1	Chinese Academy of Sciences	CHN	51	1,051	20	20.61	3	31	17
2	Int Bus Machines (IBM)	USA	41	4,606	17	112.34	18	22	1
3	CNRS	FRA	36	477	13	12.89	1	24	12
4	Harbin Institute of Technology	CHN	33	512	12	15.52	0	24	9
5	Deakin U	AUS	29	306	11	10.55	1	20	8
6	U Illinois Chicago	USA	27	4,032	14	149.33	2	20	5
7	Tsinghua U	CHN	26	483	10	18.58	3	15	8
8	Simon Fraser U	CAN	24	818	12	34.08	6	15	3
9	Chinese U Hong Kong	CHN	24	435	12	18.13	4	16	4
10	Zhejiang U	CHN	24	360	11	15	1	12	11
11	Carnegie Mellon U	USA	23	668	11	29.04	3	17	3
12	Nanyang Technological U	SGP	21	469	12	22.33	6	13	2
13	U Technology Sydney	AUS	19	514	9	27.05	3	12	4
14	Monash U	AUS	19	397	11	20.90	1	9	9
15	U Melbourne	AUS	18	604	9	33.56	2	10	6
16	Nanjing U	CHN	17	3,995	8	235	2	7	8
17	Beihang U	CHN	17	421	9	24.76	1	8	8
18	U California Riverside	USA	16	1,938	10	121.13	6	7	3
19	U Granada	SPA	16	1,424	12	89	0	15	1
20	Rutgers U New Brunswick	USA	16	363	10	22.69	5	10	1
21	National U Defense Technology	CHN	16	240	8	15	0	6	10
22	National Yang Ming Chiao Tung U	TWN	15	496	10	33.07	2	13	0
23	U Illinois Urbana Champaign	USA	15	347	8	23.8	5	5	5
24	Peking U	CHN	15	274	8	18.27	5	9	1
25	U Waterloo	CAN	15	267	9	17.8	3	12	0
26	National Cheng Kung U	TWN	15	247	7	16.47	0	13	2
27	U Science Technology CAS	CHN	15	231	10	15.4	0	9	6
28	Microsoft	USA	14	495	9	35.36	2	9	3
29	Purdue U	USA	14	273	8	19.5	3	11	0
30	Fudan U	CHN	14	240	7	17.14	2	4	8

Abbreviations are available in Table 7

submission volume. This expansion reflects not only the journal's higher impact and reputation but also the globalization of data-driven research, where intelligent systems and knowledge-based methodologies have become central across disciplines. As KAIS continues to attract diverse authors from an expanding set of countries and regions, it is anticipated that its geographical footprint and intellectual diversity will further strengthen, reinforcing its status as a truly global forum for knowledge and information systems research.

Taken together, the country-level and temporal analyses presented in Tables 10 and 11 illustrate the progressive internationalization and diversification of KAIS. The data reveal both spatial and temporal dimensions of the journal's evolution: while the United States

Table 10 The most productive and influential countries in KAIS

R	Country	TP	TC	H	C/P	D1	D2	D3	P/Po	C/Po
1	PR China	599	13,516	46	22.56	40	274	285	0.43	9.60
2	USA	502	18,272	55	36.40	124	292	85	1.48	53.74
3	India	174	2,107	24	12.11	3	62	109	0.12	1.45
4	Australia	150	6,803	29	45.35	19	76	55	5.56	251.96
5	Spain	130	3,667	27	28.21	0	95	35	2.66	75.14
6	Canada	109	2,570	27	23.58	23	69	17	2.64	62.23
7	France	105	1,504	19	14.32	7	65	33	1.53	21.96
8	UK	91	5,207	23	57.22	18	45	28	1.32	75.25
9	Italy	87	1,021	18	11.75	12	38	37	1.48	17.31
10	Germany	85	1,349	17	15.87	14	42	29	1.02	16.15
11	Taiwan	72	1,546	22	21.47	8	53	11	3.08	21.47
12	Iran	72	907	16	12.60	1	36	35	0.79	9.90
13	Japan	64	4,298	15	67.15	9	36	19	0.52	34.66
14	Brazil	57	772	14	13.54	2	18	37	0.27	3.64
15	Greece	54	1,014	16	18.78	7	29	18	5.19	97.40
16	South Korea	48	407	12	8.48	5	21	22	0.93	7.90
17	Singapore	47	965	17	20.53	13	23	11	7.80	180.83
18	Pakistan	43	631	15	14.67	0	20	23	0.17	2.51
19	Turkey	39	481	12	12.33	0	24	15	0.46	5.63
20	Saudi Arabia	33	547	12	16.58	0	23	10	0.93	15.5
21	Poland	26	485	11	18.65	0	15	11	0.71	13.3
22	Netherlands	25	1,321	12	52.84	8	10	7	1.39	73.39
23	Malaysia	25	636	11	25.44	2	15	8	0.70	17.87
24	Belgium	25	396	11	15.84	2	14	9	2.10	33.28
25	Finland	25	301	11	12.04	6	9	10	4.46	53.75
26	Israel	24	688	11	28.67	5	14	5	2.40	68.8
27	Portugal	22	445	9	20.23	1	13	8	2.06	41.59
28	Tunisia	22	314	10	14.27	1	12	9	1.79	25.53
29	Sweden	19	253	9	13.32	0	12	7	1.79	23.87
30	Mexico	16	243	10	15.19	5	7	4	0.12	1.86

Abbreviations are available in Table 7 except for: P/Po and C/Po = Papers and cites per million inhabitants

initially dominated its authorship base, China has emerged as the principal driver of recent growth, with India and several European and Asia-Pacific nations also strengthening their presence. This shift reflects the global diffusion of expertise in artificial intelligence, data science, and intelligent decision-support systems—the thematic pillars of KAIS. The sustained rise in multi-country contributions signals that the journal is evolving from a primarily Western outlet into a truly global platform for interdisciplinary research, reinforcing its role as a catalyst for knowledge integration and cross-regional scientific collaboration.

Figure 4 visualizes the global collaboration network among countries contributing to KAIS, generated using the Bibliometrix software. The map highlights China's central role

Table 11 Annual number of papers classified by countries (2004–2024)

R	Country	Total	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	PR China	599	1	1	8	5	14	11	19	18	18	25	24	30	31	31	25	53	41	45	40	71	88
2	USA	502	12	21	18	20	27	26	12	30	23	24	30	30	34	30	23	31	25	20	18	31	17
3	India	174	1	1	0	0	1	0	1	2	6	2	5	5	7	8	10	16	8	5	13	25	58
4	Australia	150	1	3	2	2	7	4	3	3	3	12	4	6	12	11	7	15	19	10	5	12	9
5	Spain	130	0	0	0	0	0	0	5	7	11	10	13	9	5	7	12	16	9	7	4	7	8
6	Canada	109	4	1	4	7	4	3	9	11	5	11	6	2	10	6	5	4	6	1	3	3	4
7	France	105	0	3	1	2	0	1	2	2	8	3	6	10	8	8	9	9	8	7	5	7	6
8	UK	91	2	4	5	3	2	2	6	4	3	3	5	3	3	7	6	5	5	6	5	2	10
9	Italy	87	1	3	1	1	2	4	1	6	4	5	3	2	5	5	1	6	7	3	6	9	12
10	Germany	85	4	0	2	1	3	4	4	4	4	5	2	4	3	5	6	5	8	2	8	8	3
11	Taiwan	72	1	2	0	1	2	2	3	3	4	8	7	2	7	3	6	10	5	2	2	0	2
12	Iran	72	0	0	0	0	0	1	1	2	2	1	2	7	3	4	3	11	9	4	1	5	16
13	Japan	64	1	2	3	1	1	1	2	7	4	7	3	0	4	3	0	6	6	4	0	3	6
14	Brazil	57	0	0	0	1	0	1	1	3	0	3	0	1	1	1	3	5	7	5	5	7	13
15	Greece	54	0	1	2	1	0	3	3	2	3	3	2	2	5	2	2	5	3	3	4	1	7
16	South Korea	48	2	0	0	0	0	3	2	2	3	2	5	2	1	2	2	0	6	4	4	3	5
17	Singapore	47	1	2	1	3	4	2	1	3	3	4	3	2	1	3	2	1	4	0	4	3	0
18	Pakistan	43	0	0	0	0	0	0	2	0	3	2	3	2	2	2	0	4	0	3	7	4	9
19	Turkey	39	0	0	0	0	0	0	0	0	3	0	1	5	4	3	4	4	0	2	0	6	7
20	Saudi Arabia	33	0	0	0	0	0	0	0	0	0	3	1	4	3	4	2	6	2	1	1	2	4
21	Poland	26	0	0	0	0	0	0	2	1	2	2	1	0	2	2	1	2	3	2	2	2	2
22	Netherlands	25	3	1	0	2	1	1	1	2	1	1	0	2	1	0	1	1	4	0	1	1	1

Table 11 (continued)

R	Country	Total	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
23	Malaysia	25	0	0	0	0	1	1	0	0	2	0	0	0	2	2	2	7	1	0	3	0	4
24	Belgium	25	1	0	0	0	0	1	1	0	3	0	3	2	0	3	1	1	3	2	1	1	2
25	Finland	25	0	2	0	1	3	0	2	1	2	0	1	2	0	0	0	1	3	1	4	2	0

R = Rank; 04–24 = Annual number of documents published between 2004–2024

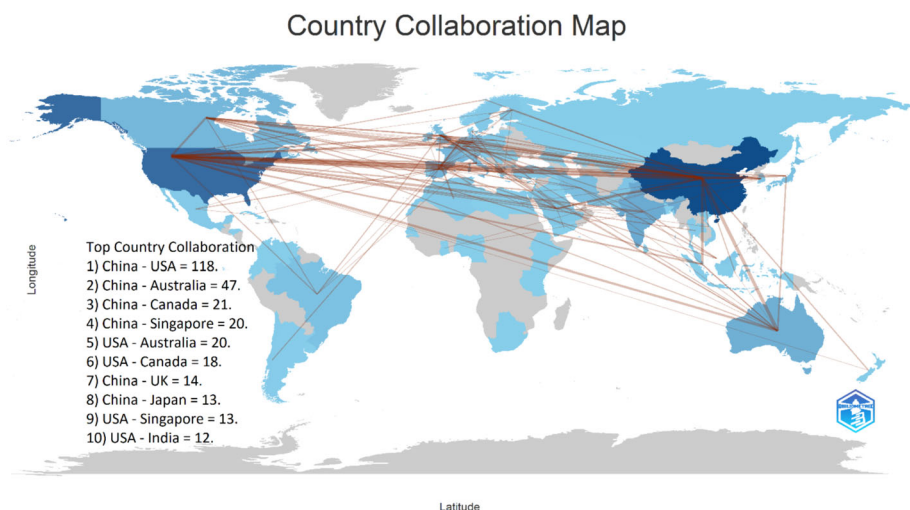


Fig. 4 Country collaboration map of KAIS

in the international research landscape, as it maintains strong collaborative links with the United States, Australia, Canada, and Singapore. Additionally, robust cooperation patterns are observed among the United States, Australia, and Canada, reflecting well-established research alliances in data-driven and computational intelligence fields. These collaboration clusters demonstrate KAIS's success in fostering international scientific exchange and cross-institutional knowledge integration.

To complement this network analysis, Table 12 provides a regional overview of KAIS publications. The results indicate that Asia is currently the most prolific region, followed by Europe and North America. Despite this, North America leads in total citations and maintains the highest h-index, underscoring the long-term influence and citation longevity of its research output. Oceania exhibits remarkable efficiency, ranking highest in both publications per million inhabitants and citations per million inhabitants, a performance that demonstrates the high research impact achieved by a region with relatively smaller population and publication volume.

Table 12 Publication structure classified by supranational regions

R	Region	TP	TC	H	C/P	D1	D2	D3	P/Pop	C/Pop
1	Asia	1,143	19,468	59	17.03	83	546	514	0.24	4.08
2	Europe	697	18,573	48	26.64	80	385	232	0.94	25.04
3	North America	593	20,443	60	34.47	142	326	125	0.98	33.72
4	Oceania	161	6,879	29	42.72	22	80	59	3.58	152.99
5	Latin America	90	1,561	19	17.34	8	38	44	0.14	2.35
6	Africa	70	842	16	12.02	2	25	43	0.05	0.56

Abbreviations are available in Tables 3 and 9

In contrast, Latin America and Africa display lower overall productivity but maintain a steady pattern of contributions, indicating the gradual expansion of the journal's global authorship base. From a broader perspective, these results confirm that KAIS has evolved into a truly global journal, encompassing active participation from all continents. The observed collaboration dynamics—anchored by strong Asia–Pacific and transatlantic partnerships—reflect the increasingly internationalized and networked nature of knowledge production in computer and information sciences.

5 Mapping KAIS with VOSviewer and Bibliometrix software

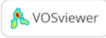
This section provides a visual exploration of bibliographic structures to enhance the understanding of the intellectual organization and dynamics within the KAIS literature. The analysis employs three complementary network perspectives: co-citation networks [87], which reveal intellectual linkages through shared citations, bibliographic coupling networks [51], which capture research relatedness based on shared references, and keyword co-occurrence networks [15], which illustrate thematic associations and emerging research fronts.

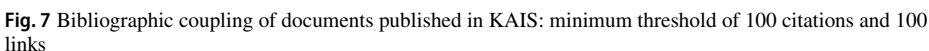
All visualizations were generated using VOSviewer [94] and the Bibliometrix package via the Biblioshiny web interface [5]. Unless otherwise specified, the figures are based on WoS data covering the period 2004–2024, providing a comprehensive and longitudinal view of the journal's bibliometric landscape.

5.1 Co-citation analysis of KAIS

Co-citation [87] as mentioned earlier, refers to a situation in which a third document cites two other documents within KAIS. The software VOSviewer creates visual bibliographic maps based on co-citation. Figure 5 showcases the co-citation network of journals referenced in KAIS. This visualisation is generated by establishing a minimum citation threshold of 100 and showing the 100 most frequent co-citation links. Each node represents a journal, with its size indicating the total number of citations it has received. The connections between the journals are shown by lines, where a thicker line suggests a higher frequency of joint co-citations. Additionally, the closeness of the nodes indicates a stronger relationship between the journals. Each colour represents a distinct cluster. Note that all the figures presented in this section with VOSviewer (Figs. 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15) use an algorithm with four units of attraction and one unit of repulsion.

The co-citation network reveals that KAIS is the most frequently cited journal, followed by *Lecture Notes in Computer Science*, *IEEE Transactions on Knowledge and Data Engineering*, *Lecture Notes in Artificial Intelligence*, and *Arxiv*. As illustrated in Fig. 5, the network organizes into five principal clusters—blue, red, purple, green, and yellow—each representing a distinct yet interconnected area within the broader field of computer and information sciences. The blue cluster encompasses KAIS and *Lecture Notes in Computer Science*, forming the core domain of knowledge systems, advanced information systems, and computational frameworks. The red cluster brings together journals such as *Information Sciences* and *Expert Systems with Applications*, emphasizing knowledge engineering, intelligent systems, and decision-support technologies. The purple cluster, anchored by *Lecture Notes in Artificial Intelligence* and *Artificial Intelligence*, reflects the conceptual and algorithmic





The green cluster focuses on machine learning and data mining, represented by outlets such as the *Journal of Machine Learning Research* and the *IEEE International Conference on Data Mining*. This group highlights the journal's strong methodological connection to computational intelligence and predictive analytics. Finally, the yellow cluster encompasses computing and information science, including *Information Processing & Management* and *Bioinformatics*, illustrating KAIS's expanding engagement with data-intensive applications across diverse scientific domains.

Table 13 lists the twenty-five most cited journals in KAIS across three intervals: 2004–2014, 2015–2019, and 2020–2024. The objective is to see how the main influences of the journal have evolved through time. *Knowledge and Information Systems* leads with 3,124 citations, followed by *Lecture Notes in Computer Science* at 2,708 and *IEEE Transactions on Knowledge and Data Engineering* at 1,696 citations. The *Lecture Notes in Artificial Intelligence* is fourth with 1,261 citations, and *Arxiv* is fifth with 1,212 citations.

Table 13 Most cited journals in KAIS: Global and temporal analysis

R	Global		2020–2024		2015–2019		2004–2014	
	Journal	Cit	Journal	Cit	Journal	Cit	Journal	Cit
1	Knowl Inf Syst	3,124	Arxiv	1,193	Lect Notes Comput Sc	842	Knowl Inf Syst	2,029
2	Lect Notes Comput Sc	2,708	Lect Notes Comput Sc	1,035	IEEE T Knowl Data En	565	Lect Notes Comput Sc	831
3	IEEE T Knowl Data En	1,696	IEEE T Knowl Data En	670	Knowl Inf Syst	471	Lect Notes Artif Int	491
4	Lect Notes Artif Int	1,261	Expert Syst Appl	663	Lect Notes Artif Int	372	IEEE T Knowl Data En	461
5	Arxiv	1,212	Knowl Inf Syst	624	J Mach Learn Res	312	Proc Int Conf Data	342
6	Expert Syst Appl	1,095	AAAI Conf Artif Inte	600	Expert Syst Appl	297	Mach Learn	316
7	J Mach Learn Res	923	Adv Neur In	538	Proc Int Conf Data	291	SIGMOD Record	231
8	Inform Sciences	882	IEEE Access	535	Proc VLDB Endow	270	Data Min Knowl Disc	216
9	Proc Int Conf Data	880	Inform Sciences	494	IEEE Data Mining	257	J Mach Learn Res	208
10	Mach Learn	717	Knowl-Based Syst	468	Inform Sciences	248	IEEE T Pattern Anal	175
11	IEEE Data Mining	712	Proc CVPR IEEE	460	Data Min Knowl Disc	223	IEEE Data Mining	162
12	Adv Neur In	698	J Mach Learn Res	403	IEEE T Pattern Anal	223	Commun ACM	153
13	Knowl-Based Syst	689	Lect Notes Artif Int	398	Mach Learn	221	Inform Sciences	140
14	Data Min Knowl Disc	676	Pr Mach Learn Res	322	Knowl-Based Syst	178	Bioinformatics	138
15	Proc VLDB Endow	663	Proc VLDB Endow	322	Pattern Recogn	175	Expert Syst Appl	135
16	AAAI Conf Artif Inte	656	Neurocomputing	301	Commun ACM	145	Pattern Recogn	131
17	IEEE T Pattern Anal	650	IEEE Data Mining	293	Neurocomputing	135	Artif Intell	116
18	Proc CVPR IEEE	589	IEEE T Pattern Anal	252	ACM Comput Surv	126	ACM T Database Syst	114
19	Pattern Recogn	556	Pattern Recogn	250	P Natl Acad Sci USA	113	P ACM SIGMOD Int	112
							C M	

Table 13 (continued)

R	Global		2020–2024		2015–2019		2004–2014	
	Journal	Cit	Journal	Cit	Journal	Cit	Journal	Cit
20	IEEE Access	550	Proc Int Conf Data	247	Artif Intell	109	Inform Process Manag	109
21	Neurocomputing	486	Data Min Knowl Disc	237	Thesis	108	Data Knowl Eng	107
22	ACM Comput Surv	439	ACM Comput Surv	227	SIGMOD Record	105	Fuzzy Set Syst	103
23	Commun ACM	431	Appl Soft Comput	211	Data Knowl Eng	102	SIAM Proc S	94
24	SIGMOD Record	407	Mach Learn	180	Phys Rev E	98	IEEE T Neural Networ	88
25	Pr Mach Learn Res	356	Appl Intell	179	VLDB J	97	ACM Comput Surv	86

R = Rank; Cit = Citations

During the first period (2004–2014), KAIS emerged as the primary citation source, reflecting the journal's self-referential consolidation phase, when KAIS began to establish its identity as a central venue for research in knowledge discovery and intelligent systems.

In the second period (2015–2019), *Lecture Notes in Computer Science* took the lead, signalling a strong methodological exchange between KAIS and computational research communities focused on data-driven learning and algorithmic modeling. *IEEE Transactions on Knowledge and Data Engineering* maintained a consistent third position across all three intervals, highlighting its long-standing methodological alignment with KAIS. A notable trend in the most recent period (2020–2024) is the rapid ascent of *Arxiv*, which now occupies a dominant position in contemporary citations. This shift underscores the growing influence of open-access and preprint dissemination in shaping scholarly communication. The increased referencing of Arxiv papers indicates that KAIS authors are increasingly engaging with cutting-edge, pre-publication research, reflecting the journal's responsiveness to the accelerated pace of innovation in computer science and artificial intelligence.

In short, these temporal transformations reveal that KAIS's citation framework has transitioned from an insular, foundational stage to a globally interconnected, preprint-centric milieu, paralleling the broader evolution of knowledge dissemination in the digital era.

Figure 6 depicts the document co-citation network within KAIS, corresponding to the results presented in Table 5. The visualization applies a minimum citation threshold of 20 and includes the 100 strongest co-citation links. Each cluster represents a coherent intellectual community shaped by shared methodological or conceptual orientations.

The red cluster dominates the network, anchored by two seminal contributions: Demšar [24], which established robust statistical foundations for the comparison of classifiers, and Wu et al. [98], which codified the “Top 10 Algorithms in Data Mining”. Together, these works form the methodological nucleus of KAIS research, representing the journal's long-standing emphasis on rigorous evaluation frameworks and benchmark-setting contributions that unify the data mining community.

The blue cluster highlights the neural and representation learning revolution, led by Vaswani et al. [96] on attention mechanisms and Mikolov et al. [71] on word embeddings. Their prominence signals KAIS's integration with contemporary deep learning paradigms, reflecting a disciplinary shift from classical rule-based approaches to representation-driven intelligence.

The yellow cluster, centred around Agrawal et al. [2] and Agrawal & Srikant [1], embodies the origins of association rule mining and frequent itemset discovery. These foundational studies established the algorithmic and conceptual underpinnings of pattern discovery, providing a historical anchor that continues to inform KAIS's focus on knowledge extraction and data-driven reasoning.

Finally, the green cluster features key contributions from Blei et al. [12] on Latent Dirichlet Allocation and Jain et al. [45] on clustering analysis. This group bridges probabilistic modeling and unsupervised learning, illustrating the journal's engagement with topics that lie at the intersection of machine learning, information retrieval, and knowledge representation.

In summary, the structure of this co-citation network reveals KAIS's dual identity: on one hand, it preserves the foundational traditions of data mining and statistical learning, while on the other, it actively assimilates emerging paradigms in deep learning and probabilistic modeling. This intellectual balance reinforces the journal's position as a bridge between classical data-driven analytics and modern AI methodologies, contributing to the continuous renewal of knowledge and information systems research.

5.2 Bibliographic coupling of KAIS

Bibliographic coupling occurs when two documents share common references, indicating that they address closely related research themes [51]. Figure 7 visualizes the bibliographic coupling network of documents published in KAIS, applying a minimum threshold of 100 for both document citations and link strength. The size of each node corresponds to the total number of citations received in the WoS, while the proximity and density of links illustrate the degree of intellectual similarity among the journal's publications. This visualization complements Table 4, as the most highly cited documents—Wu et al. [98], Keogh and Ratanamahatana [49], and Štrumbelj and Kononenko [88]—occupy central positions within the network.

From an interpretive perspective, the dominant nodes reveal KAIS's methodological coherence around several enduring and emerging intellectual pillars.

First, Wu et al. [98] defines the benchmarking foundation of data mining algorithms, shaping how subsequent KAIS works structure methodological comparison and evaluation standards. Second, Keogh and Ratanamahatana [49] anchors the time-series analytics tradition, linking to a dense subnetwork of studies on temporal modeling, similarity search, and streaming data. Third, Štrumbelj and Kononenko [88] exemplifies the journal's growing engagement with explainable and interpretable machine learning, bridging classical data mining with modern frameworks of model transparency and trustworthiness.

Recent additions have further diversified and modernized the bibliographic landscape. The paper by Hu et al. [42] on Model Complexity of Deep Learning connects theoretical and empirical research streams by formalizing the relationship between model structure, optimization behaviour, and generalization performance—thereby extending KAIS's methodological base into the theoretical foundations of deep neural systems. Similarly, Kirişçi [53] introduces innovative cosine similarity and distance measures for Fermatean fuzzy sets, expanding the journal's scope toward decision intelligence and uncertainty modeling, and reinforcing KAIS's reputation for methodological rigor and cross-domain applicability.

The clustering of these classical and contemporary works demonstrates that KAIS has developed a densely interconnected research corpus, where foundational methodological papers coexist with cutting-edge research in explainable, distributed, and fuzzy-based intelligent systems. This structure highlights KAIS's dual role as both a repository of canonical methods and a platform for continual methodological innovation—a defining characteristic of its sustained and evolving scholarly impact within the fields of data mining, artificial intelligence, and information systems.

Figure 8 presents the bibliographic coupling network of KAIS authors with at least five publications and a minimum of 100 connections. The cluster colours indicate the average publication year, providing a temporal dimension to author activity. Researchers shown in purple and blue correspond to early contributors (2014–2016), whereas those in orange and red represent more recent authors who have published primarily over the last four years. Node size reflects the relative citation impact of each author's work, showing comparable prominence among the leading contributors.

The network exhibits a moderately dispersed yet cohesive structure, reflecting a diverse and evolving research community within KAIS. The purple cluster highlights established figures such as Philip S. Yu and Jian Pei, whose influential contributions have solidified the journal's methodological foundation in data mining and knowledge discovery. For instance, Yu's work in Gedik et al. [35] introduced an efficient CPU load-shedding mechanism for binary stream joins, addressing scalability challenges in continuous query processing and laying the groundwork for modern stream data management systems. Similarly, Pei's research

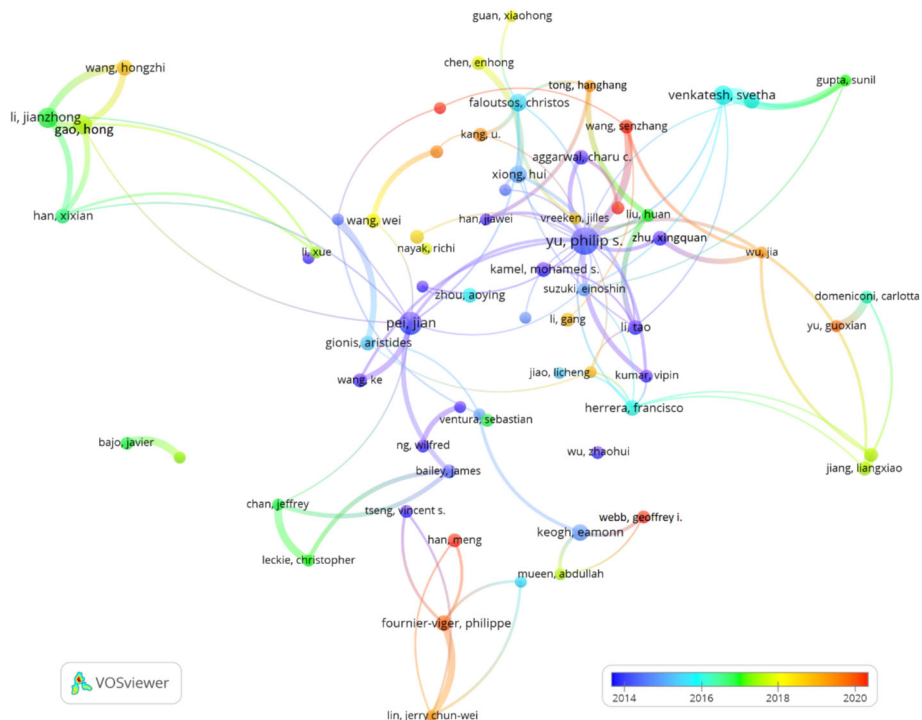


Fig. 8 Bibliographic coupling of authors publishing in KAIS: minimum publication threshold of 5 documents and 100 links

in Tang et al. [97] provided one of the earliest comprehensive studies on email mining, outlining its tasks, techniques, and tools, and broadening the application scope of text and knowledge mining in real-world communication systems.

The blue cluster, featuring Christos Faloutsos and Svetha Venkatesh, highlights major contributions to graph analytics and temporal data mining, underscoring the integration of network science and intelligent systems into KAIS's research agenda. For instance, Faloutsos's co-authored paper in Jiang et al. [46] introduced a novel framework for inferring lock-step behaviour from connectivity patterns in large-scale graphs, enabling the detection of coordinated or anomalous group activities in complex networks such as social media and cybersecurity systems. Likewise, Venkatesh's contribution in Saha et al. [82] proposed an effective sparse imputation method for electronic medical records, improving emergency risk prediction in healthcare through temporal modeling and probabilistic reasoning.

The green cluster, featuring Jianzhong Li and Liangxiao Jiang, underscores KAIS's expanding focus on large-scale data processing, sensor networks, and intelligent information systems. In Zheng et al. [106], Li contributed to a comprehensive survey on query result diversification, providing a unified framework that balances relevance and diversity in information retrieval and recommendation systems—an essential issue for real-time data environments. Similarly, Li et al. [62], co-authored by Jiang, proposed an enhanced value difference metric with attribute weighting, improving the discriminative power of classification algorithms in heterogeneous datasets.

Finally, the red cluster, which includes Philippe Fournier-Viger and Geoff Webb, represents the most recent generation of influential authors contributing to KAIS's evolving research frontier. Fournier-Viger, through Duong et al. [28], advances the journal's long-standing tradition in high-utility pattern mining by proposing CG-FHAUI, an algorithm designed for succinct pattern discovery that enhances both computational efficiency and interpretability in large-scale transactional data. In parallel, Webb, in Lee et al. [61], introduces a novel framework for computing marginal and conditional divergences between decomposable models, enabling precise statistical model comparison in high-dimensional contexts, including quantum computing and earth observation. Overall, the network underscores KAIS's evolution toward interpretable, distributed, and AI-driven systems, consolidating its role as a leading forum for methodological advancement.

Figure 9 depicts the bibliographic coupling network among institutions that have published at least five papers in KAIS. This figure aligns with the results shown in Table 9 and offers a visual overview of the institutional relationships underlying the journal's scholarly ecosystem. The network reveals that stronger connections typically occur among institutions within the same country, reflecting the persistence of national collaboration clusters. For instance, the thick link between the Chinese Academy of Sciences and Nanjing University illustrates the strength of domestic co-authorship within China's research system. The visualization is

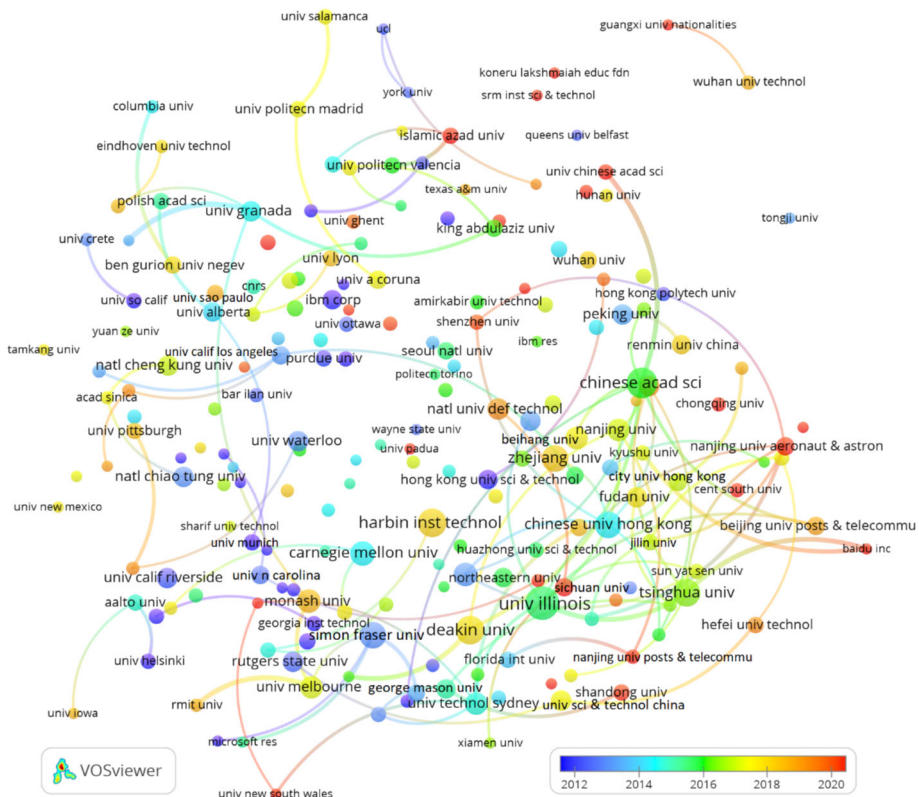


Fig. 9 Bibliographic coupling of institutions publishing in KAIS: minimum publication threshold of 5 documents and 100 links

inherently shaped by multi-authored works, since co-authorship entails shared bibliographic references, reinforcing inter-institutional ties.

The colour gradient indicates the temporal distribution of institutional activity: institutions in purple and blue were most active during 2012–2014, representing KAIS's early contributors, whereas those in orange and red correspond to recently emerging institutions, publishing mainly over the last four years. The comparable node sizes suggest a relatively balanced level of institutional influence, though a notable increase in interconnections between institutions is observed from 2020 onward, marking an acceleration in global cooperation and cross-institutional research intensity.

The cluster composition provides additional insight into KAIS's collaborative landscape. The purple cluster includes Simon Fraser University (Canada) and the University of California, Riverside (USA), which played a foundational role during the journal's formative years. The blue cluster highlights the Chinese University of Hong Kong (China) and Carnegie Mellon University (USA), symbolizing early trans-Pacific collaboration. In the green cluster, the Chinese Academy of Sciences (China) and the University of Illinois at Chicago (USA) emerge as central hubs with sustained productivity. The orange cluster, dominated by the Harbin Institute of Technology (China) and Deakin University (Australia), represents expanding cooperation between Asia and Oceania. Finally, the red cluster underscores recent growth driven by Monash University (Australia) and Nanjing University (China), whose increasing co-publication intensity reflects the journal's contemporary research dynamics.

Figure 10 extends this analysis to the country level, using a minimum publication threshold of five documents and 50 bibliographic links. Node sizes correspond to the number of KAIS publications, consistent with Table 10. The results reaffirm China's leading position, followed by the United States, India, Australia, and Spain. Five major country clusters emerge—purple, blue, green, yellow, and red—each representing distinctive collaboration patterns and temporal trajectories. The purple cluster, led by the USA and Canada, indicates strong North American cooperation. The blue cluster, featuring the UK and Germany, captures Europe's sustained engagement in the field. The green cluster, anchored by Australia and Spain, reflects long-standing ties between Oceania and Europe. The yellow cluster is centred on China and Saudi Arabia, highlighting Asia's expanding scientific influence. Finally, the red cluster, comprising India and Brazil, represents a new generation of rapidly emerging contributors.

Taken together, these networks demonstrate KAIS's progressive internationalization and the emergence of a globally distributed research community. The journal's collaborative landscape has evolved from regional concentrations into a complex, interconnected system of cross-national partnerships, reinforcing its position as a leading international platform for interdisciplinary research in knowledge and information systems.

In summary, the bibliographic coupling analyses—spanning documents, authors, institutions, and countries—reveal a well-integrated and progressively global research network within KAIS. The findings demonstrate how intellectual cohesion, regional collaboration, and methodological continuity have converged to strengthen the journal's international visibility. Building upon this structural understanding, the next section explores keyword co-occurrence networks, providing a complementary perspective on the conceptual and thematic evolution of KAIS research over the past two decades.

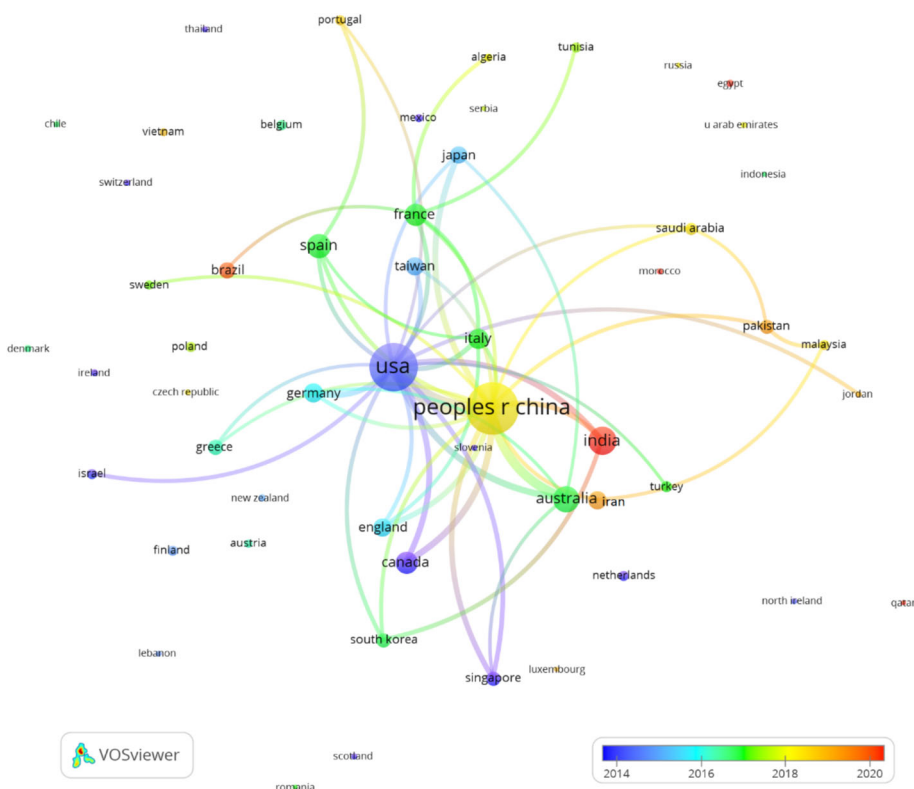


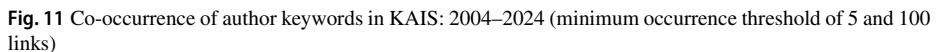
Fig. 10 Bibliographic coupling of countries publishing in KAIS: minimum publication threshold of 5 documents and 50 links

6 Keyword and topic analysis of KAIS

Co-occurrence of keywords refers to the simultaneous appearance of two or more terms within the same publication [15]. Examining keyword co-occurrence is a powerful bibliometric approach for uncovering the conceptual structure, emerging themes, and evolutionary trends of a journal [30].

The co-occurrence network of author keywords (Fig. 11) depicts the interrelationships among terms used in KAIS publications between 2004 and 2024, applying a minimum occurrence threshold of five and a maximum of 100 links. In this visualization, each node represents a keyword, with node size proportional to its frequency of occurrence, and link thickness reflecting co-occurrence strength. Colours indicate the average year of appearance of each keyword, thus revealing the temporal evolution of research topics in KAIS.

The network highlights “data mining” and “classification” as the dominant and most central terms, forming the core conceptual nucleus of the journal. These keywords encapsulate KAIS’s enduring focus on methodological rigor and knowledge discovery, serving as the foundation upon which subsequent advances in intelligent systems have been constructed. Building upon this foundation, Li et al. [63] extend the journal’s intellectual trajectory by integrating data mining, AI ethics, and human–AI collaboration within a unified interpretability



Closely related, “clustering” and “machine learning” function as interdisciplinary bridges connecting algorithmic design with practical knowledge applications. Liu et al. [65] exemplify this orientation, reinforcing KAIS’s role as a nexus between theoretical learning architectures and real-world socio-technical systems, especially in the domains of federated and privacy-preserving learning.

Together, these developments demonstrate KAIS's intellectual evolution from classical data mining toward an integrated, ethically aware, and theoretically grounded AI ecosystem, positioning the journal as a leading venue at the intersection of methodological innovation and responsible knowledge systems.

 Springer

Beyond simple frequency analysis, the spatial and temporal patterns in the network reveal a progressive thematic convergence between classical data mining and modern artificial intelligence. Earlier clusters (2004–2012) were dominated by topics such as association rules, fuzzy systems, and decision support, whereas more recent years (post-2016) show an increasing prevalence of deep learning, graph analytics, and explainable AI. This shift demonstrates KAIS’s ability to evolve conceptually with technological advances while maintaining a strong foundation in knowledge discovery and intelligent systems.

Figure 12 depicts the co-occurrence network of author keywords for publications originating from North American institutions, applying a minimum threshold of three keyword



Fig. 12 Co-occurrence of author keywords in KAIS: North America (minimum occurrence threshold of 3 and 100 links)

occurrences and visualizing the top 100 co-occurrence links. The colour gradient indicates the average publication year of each keyword, allowing the temporal evolution of research topics to be observed.

The analysis reveals two dominant thematic clusters. The purple cluster centred around “data mining”, “clustering”, and “outlier detection”, reflects North America’s historical emphasis on core data analysis and pattern recognition techniques—fields that established the conceptual foundations of KAIS and remain influential within the journal’s methodological framework. The blue cluster, comprising “classification” and “privacy”, represents an early convergence between machine learning and data ethics, signalling the region’s pioneering role in trustworthy and privacy-preserving data mining. The temporal overlay shows a strong interconnection between these clusters during the 2012–2014 period, corresponding to the maturation of data-intensive learning and the emergence of privacy-aware computing.

In more recent years, newer clusters—represented in orange and red hues—highlight the rise of “deep learning”, “graph neural networks”, and “entity linking” as key topics. These indicate North America’s transition toward representation learning, semantic modeling, and knowledge integration, areas that align closely with contemporary artificial intelligence research. The coexistence of traditional data mining concepts with cutting-edge deep learning themes illustrates a continuum of methodological evolution, rather than a complete paradigm shift. From a broader perspective, the figure also suggests that North American publications tend to be older on average than those from other regions, reflecting the historical leadership of North American institutions in the early development of knowledge systems and data mining. Although temporal differences are not substantial, they reveal that North America’s contributions provided much of the methodological groundwork upon which later regional innovations have built.

Overall, the keyword co-occurrence map for North America underscores the region’s dual legacy: it has anchored the foundational methods of data science while simultaneously embracing emerging paradigms in deep and graph-based learning, maintaining a leading role in shaping the intellectual trajectory of KAIS.

Figure 13 illustrates the co-occurrence of author keywords in KAIS for Europe, with a minimum occurrence threshold of three and showing the 100 strongest co-occurrence links.

The network reveals a well-defined conceptual organization that reflects the region’s balanced emphasis on both methodological innovation and applied research. The blue cluster, anchored by “data mining”, “classification”, and “clustering”, represents Europe’s sustained leadership in algorithmic design, knowledge discovery, and pattern recognition. This cluster captures Europe’s contribution to establishing empirical rigor and benchmarking standards within the KAIS community.

The green cluster, centred on “multi-agent systems” and “concept drift”, highlights Europe’s pioneering role in adaptive and distributed intelligence. European researchers have been instrumental in advancing frameworks that model dynamic environments, autonomous decision-making, and temporal learning, key components of intelligent system design. The yellow cluster, featuring “machine learning”, “feature selection”, and “graph mining”, signifies a transition toward data-driven knowledge modeling and network-based reasoning, which connect classical data mining with modern graph analytics.

More recent developments, represented in the red cluster, include “process mining”, “deep learning”, and “big data”—keywords that indicate Europe’s growing engagement with data-intensive automation and AI-driven process analysis. The prominence of “process mining” is particularly notable, as it underscores Europe’s distinct research identity within KAIS: bridging information systems, workflow analytics, and organizational intelligence. Overall, the European network suggests a methodologically mature ecosystem, where classical

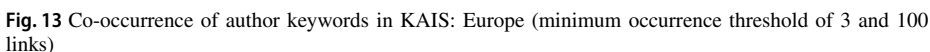


Figure 14 shows the co-occurrence of author keywords in KAIS for East Asia, using a minimum threshold of three occurrences and 100 links. The purple cluster, dominated by “data mining”, “clustering”, and “classification”, reflects East Asia’s extensive contribution to core data analytics and knowledge discovery methods, foundational elements that continue to define much of the region’s research output. The blue cluster, containing “feature selection”, “collaborative filtering”, and “transfer learning”, demonstrates a strong orientation toward personalization, recommendation systems, and knowledge transfer, key areas where East Asian research has achieved global recognition and application impact.

Taken together, the East Asian network reveals a trajectory of progressive specialization and technological advancement. The region’s growing emphasis on deep learning, knowledge

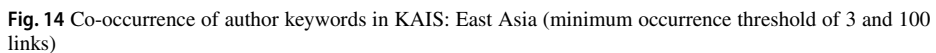
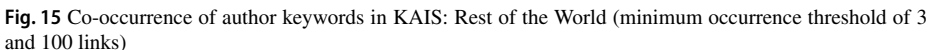


Figure 15 presents the co-occurrence network of author keywords in KAIS for the rest of the world, applying a minimum threshold of three occurrences and displaying the 100 strongest co-occurrence links. The purple cluster, centred on “data mining” and “clustering”, reflects the universal methodological foundation upon which KAIS research is built. These terms represent the journal’s historical and ongoing focus on knowledge discovery, unsupervised learning, and pattern extraction, themes that continue to define its scientific core. The green cluster, characterized by “classification” and “information retrieval”, emphasizes the interplay between machine intelligence and semantic organization, capturing the journal’s dual commitment to algorithmic rigor and knowledge structuring.

 Springer



From a comparative standpoint, the figure indicates that regional differences are relatively subtle, suggesting that the journal's readership and contributor base share a convergent conceptual vocabulary. While North American publications tend to be older—reflecting their pioneering role in establishing early methodological standards—East Asian and European contributions increasingly dominate the recent landscape, integrating deep learning, graph-based reasoning, and domain adaptation. Despite minor variations in emphasis, all regions align around a common intellectual nucleus: data mining, classification, clustering, and machine learning.

Figure 16 illustrates the evolution of trending author keywords in KAIS from 2005 to 2024, generated using Bibliometrix. Each keyword's line spans its first to third quartile of occurrence, with dot size indicating frequency and position reflecting the median year [30].

 Springer

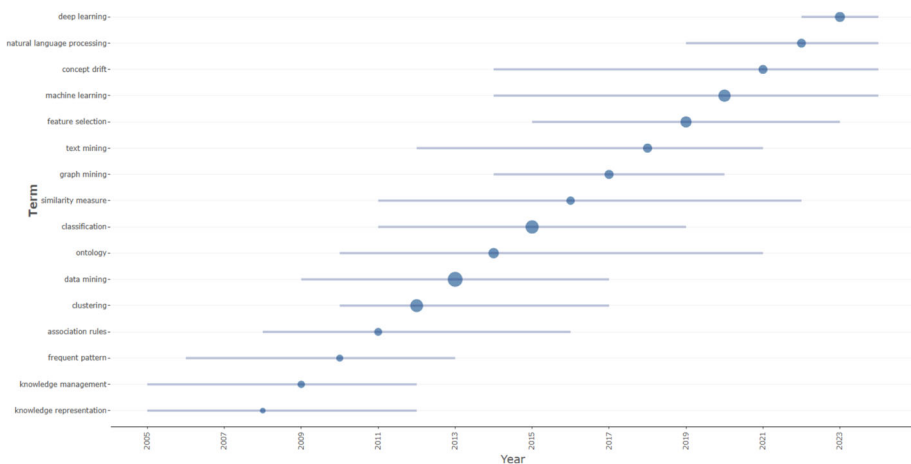


Fig. 16 Trend Topics in KAIS (2005–2024)

new themes like graph mining, feature selection, and machine learning emerged, signalling a shift toward network-based and automated learning. In the most recent period (2021–2024), deep learning, graph neural networks, and natural language processing gained prominence, reflecting the journal’s adaptation to AI-driven and adaptive systems.

Overall, the temporal pattern reveals KAIS’s transition from classical data mining toward integrated artificial intelligence, confirming its role as a central venue for the evolution of intelligent and knowledge-based systems.

Table 14 presents the 25 most frequently used author keywords in KAIS (2004–2024), grouped into three periods (2004–2014, 2015–2019, and 2020–2024). Overall, “data mining” is the most frequent keyword (150 occurrences), followed by “classification” (104), “clustering” (95), and “machine learning” (84). “Data mining” is the leading term in the first two periods, consistent with KAIS’s longstanding emphasis on knowledge discovery and pattern analysis. In the most recent period (2020–2024), however, “machine learning” and “deep learning” become more prominent, while “data mining” drops to eighth place. This change suggests a shift in terminology and thematic emphasis toward contemporary AI and representation-learning approaches, while many foundational knowledge discovery problems continue to be addressed within modern machine learning frameworks.

To avoid interpreting keywords as flat or equivalent categories, the thematic analysis is structured into hierarchical conceptual tiers. At the highest level, broad paradigmatic domains (e.g., knowledge systems, artificial intelligence, data science) define the journal’s intellectual scope. At the intermediate level, methodological families such as data mining, machine learning, and deep learning represent core analytical paradigms. At the most specific level, technical subfields—including graph neural networks, support vector machines, anomaly detection, or concept drift—correspond to specialized implementations within these broader frameworks. This hierarchical perspective enables a more meaningful interpretation of thematic evolution over time.

Thematic evolution: core, emerging, and shifting topics. To complement the static co-occurrence maps, we interpret the keyword results longitudinally using the temporal overlays

Table 14 Most frequent keywords in KAIS: Global and temporal analysis

R	Global		2020–2024		2015–2019		2004–2014	
	Keyword	Occ	Keyword	Occ	Keyword	Occ	Keyword	Occ
1	Data Mining	150	Machine Learning	46	Data Mining	45	Data Mining	87
2	Classification	104	Deep Learning	38	Classification	31	Clustering	61
3	Clustering	95	Graph Neural Network	29	Social Networks	20	Classification	48
4	Machine Learning	84	Feature Selection	28	Feature Selection	18	Ontology	27
5	Feature Selection	60	Classification	25	Clustering	17	Machine Learning	23
6	Ontology	48	Recommender Systems	21	Machine Learning	15	Information Retrieval	22
7	Social Networks	47	Social Networks	20	Big Data	14	Data Stream	16
8	Deep Learning	43	Data Mining	18	Collaborative Filtering	14	Support Vector Machines	16
9	Time Series	43	Natural Language Processing	18	Time Series	14	Feature Selection	14
10	Information Retrieval	41	Time Series	18	Recommender Systems	13	Privacy	13
11	Recommender Systems	41	Clustering	17	Transfer Learning	12	Association Rules	11
12	Collaborative Filtering	32	Ontology	16	Graph Mining	11	Knowledge Management	11
13	Data Stream	31	Knowledge Graph	15	Anomaly Detection	10	Semantic Web	11
14	Big Data	30	Big Data	14	Information Retrieval	10	Text Mining	11
15	Text Mining	30	Concept Drift	14	Multi-Agent Systems	10	Time Series	11

Table 14 (continued)

R	Global		2020–2024		2015–2019		2004–2014	
	Keyword	Occ	Keyword	Occ	Keyword	Occ	Keyword	Occ
16	Transfer Learning	30	Sentiment Analysis	14	Support Vector Machines	10	Outlier Detection	10
17	Graph Neural Network	29	Transfer Learning	14	Cloud Computing	9	Dimensionality Reduction	9
18	Graph Mining	28	Collaborative Filtering	13	Privacy	9	Graph Mining	9
19	Support Vector Machine	28	Ensemble Learning	13	Data Streams	8	Information Extraction	9
20	Concept Drift	27	Attention Mechanism	11	Text Mining	8	Multi-Agent Systems	9
21	Natural Language Processing	27	Graph Embedding	11	Community Detection	7	Text Classification	9
22	Anomaly Detection	26	Representation Learning	11	Frequent Pattern Mining	7	Document Clustering	8
23	Privacy	25	Text Mining	11	MapReduce	7	Knowledge Representation	8
24	Multi-Agent Systems	22	Graph Neural Networks	10	Massive Data	7	Query Optimization	8
25	Sentiment Analysis	22	LSTM	10	6 Keywords	7	Semi-Supervised Clustering	8

Occ = Occurrences

(Figs. 11 and 16) and the period-based frequencies (Table 14). Overall, the results indicate a stable methodological core together with a gradual expansion toward more recent AI paradigms.

Core and persistent topics. Across all periods, “data mining”, “classification”, and “clustering” remain the most recurrent keywords (Table 14), indicating that predictive modelling and knowledge discovery constitute the journal’s persistent thematic backbone. Several long-running application-oriented themes—such as “ontology”, “information retrieval”, “recommender systems”, and “social networks”—also appear consistently, reflecting sustained interest in semantic information management and user-/network-oriented systems.

Emerging topics. From 2020 onward, “deep learning” becomes one of the most frequent keywords, while terms such as “graph neural network(s)”, “natural language processing”, “knowledge graph”, and “representation learning” gain visibility (Table 14; Fig. 16). This pattern suggests an increasing focus on representation learning and graph-centric modelling, consistent with the prominence of graph- and neural-network-related topics in the SciVal-based profiles (Tables 15 and 16).

Shifting emphasis. In relative terms, “data mining” decreases in rank in the most recent period (2020–2024) as “machine learning” and “deep learning” become more prominent (Table 14). Rather than indicating a thematic decline, this shift reflects a change in terminology and modelling paradigms. Many classical data mining problems are now addressed through machine learning pipelines and neural architectures, suggesting a paradigmatic transition rather than a disciplinary replacement. In this sense, deep learning and graph neural networks can be interpreted as specialized methodological branches within the broader evolution of data-driven knowledge discovery.

Topic persistence vs. novelty at scale. The SciVal-based topic profiles complement the keyword analysis by adding citation-normalized (FWCI) and momentum-based (PP) indicators. For example, network-oriented topics (e.g., community detection and social network analysis) show high citation impact (Table 15), while graph- and neural-network-related topics exhibit very high prominence percentiles (Table 15). Taken together, these results indicate that KAIS maintains stable core themes while selectively incorporating high-momentum topics aligned with current AI research trajectories.

Collectively, the longitudinal evidence supports a three-layer thematic structure in KAIS: (i) a persistent core centred on data mining and predictive modelling, (ii) stable application domains in semantic information management and recommender/information retrieval systems, and (iii) emerging AI frontiers associated with deep learning, graph learning, and language-centric methods. This interpretation provides additional context for the journal’s thematic continuity and its recent expansions.

To broaden the keyword analysis, Fig. 17 presents a thematic map that organizes the 300 most frequently used keywords in KAIS into broader conceptual categories. The map, generated using Bibliometrix and Biblioshiny [5] with the Walktrap clustering algorithm [78], highlights the three most common keywords in each cluster.

The central blue cluster represents the core knowledge base of the journal, comprising widely recurring terms such as “data mining”, “classification”, and “clustering”. These topics form the methodological foundation of KAIS and illustrate its enduring focus on knowledge discovery, pattern recognition, and information system modeling. Their centrality in the network indicates that they serve as bridging themes, connecting multiple areas across the journal’s disciplinary spectrum.

Table 15 Leading topics in KAIS between 2014 and 2023 (Scopus)

R	Topic	TP	FWCI	PP
1	Association Rules; Frequent Itemsets; Data Mining	32	1.18	95.88
2	Time Series; Neural Network; Data Mining	29	0.74	97.45
3	Recommender Systems; Collaborative Filtering; Graph Neural Network	27	1.04	99.83
4	Community Detection; Complex Networks; Social Network Analysis	21	2.5	97.76
5	Sequential Pattern Mining; Database Systems; Data Mining	20	1.22	82.08
6	Topic Modeling; Latent Dirichlet Allocation; Data Mining	19	1.12	98.65
7	Data Stream; Learning Systems; Concept Drift	19	1.05	96.87
8	Link Prediction; Heterogeneous Network; Representation Learning	17	0.35	98.61
9	Social Network; Information Diffusion; Data Mining	17	0.71	97.84
10	Points of Interest; Social Network; Data Mining	17	1.64	97.08
11	Learning Systems; Computer Vision; Domain Adaptation	16	0.76	99.51
12	Imbalanced Data; Support Vector Machine; Class Imbalance	15	1.77	99.49
13	Recommender Systems; Collaborative Filtering; E-Commerce	13	1.5	98.53
14	Decision-Making; Database Systems; Query Processing	13	0.35	75.52
15	Subgraphs; Graphic Methods; Data Mining	12	0.47	68.72
16	Clustering Algorithm; Clustered Data; Data Mining	11	1.13	97.31
17	Graph Theory; Neural Network; Graph Convolutional Network	10	0.25	99.85
18	Social Networking (Online); Sentiment Classification; Data Mining	10	2.47	99.32
–	Crowdsourcing; Learning Systems; Artificial Intelligence	10	0.75	95.31
–	7 Topics	9	–	–
–	525 Other Topics	[1, 8]	–	–

R = Rank; TP = Total papers; FWCI = Field-weighted citation impact (data from Scopus); PP = Worldwide prominent percentile (according to Scopus and FWCI)

The light green cluster, centred on “ontology”, “information retrieval”, and “big data”, reflects the journal’s role in integrating semantic technologies with data-driven methods. This area underscores the evolution of KAIS from traditional data analysis toward semantic representation, knowledge organization, and large-scale information management—a hallmark of its interdisciplinary character.

Table 16 Leading topic clusters in KAIS between 2014 and 2023 (Scopus)

R	Topic cluster	TP	FWCI	PP
1	Data Mining; Graph Neural Network; Social Network Analysis	139	0.89	95.16
2	Sentiment Analysis; Natural Language Processing; Machine Learning	115	1.07	99.67
3	Data Mining; Artificial Intelligence; Information System	105	1.01	79.13
4	Machine Learning; Data Mining; Artificial Intelligence	88	1.29	93.39
5	Data Mining; Artificial Intelligence; e-Commerce	79	1.13	85.41
6	Image Segmentation; Deep Neural Network; Object Detection	70	1.43	100
7	Data Mining; Artificial Intelligence; Machine Learning	63	0.85	57.55
8	Data Mining; Image Segmentation; Artificial Intelligence	39	0.87	38.91
9	Commerce; Machine Learning; Transport	38	0.74	90.71
10	Ontology; Semantic Web; Linked Data	35	0.76	49.24
11	Query Processing; Cloud Computing; Input/Output	35	0.46	26.35
12	Machine Learning; Artificial Intelligence; Knowledge Acquisition	34	0.84	72.66
13	Web Service; Quality of Service; Data Mining	29	1.45	48.20
14	Biometrics; Image Processing; Artificial Intelligence	23	0.52	83.32
15	Query Processing; Cloud Computing; Internet of Things	23	0.56	23.41
16	Data Structure; Pattern Matching; Computational Complexity	17	0.63	19.22
17	Particle Swarm Optimization; Mathematical Optimization; Benchmarking	16	1.51	96.59
18	Number; Multiple-Criteria Decision Analysis; Fuzzy Logic	15	2.08	90.90
19	Forestry; Artificial Intelligence; Internet of Things	15	0.63	60.23
20	Rough Set; Data Mining; Information System	15	0.61	31.00
21	Transport; Data Mining; Information System	14	1.05	94.76
22	Causal Inference; Regression Analysis; Randomized Controlled Trial	13	0.97	81.62
23	Search Engine; Case-Based Reasoning; Data Mining	13	0.42	15.56
24	Image Processing; Robot; Object Recognition	11	0.49	73.77
25	Authentication; Network Security; Cloud Computing	10	0.54	79.79
–	3 Topic Clusters	9	–	–
–	124 Other Topic Clusters	[1, 8]	–	–

R = Rank; TP = Total papers; FWCI = Field-weighted citation impact (data from Scopus); PP = Worldwide prominent percentile (according to Scopus and FWCI)

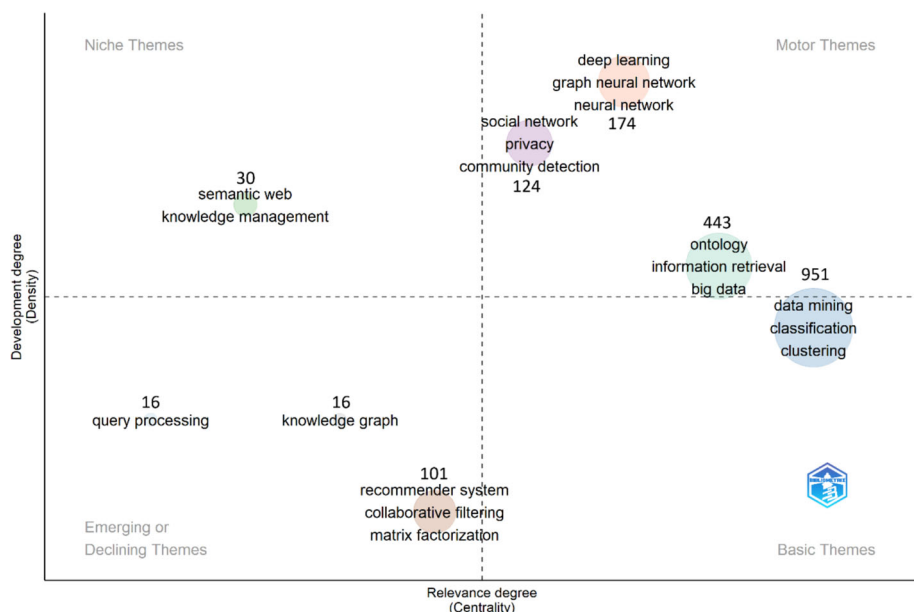


Fig. 17 Thematic map with the 300 most frequent author keywords and minimum cluster frequency of 5 (clustering algorithm: Walktrap)

The orange cluster, containing “deep learning”, “graph neural network”, and “neural network”, captures the journal’s growing engagement with modern artificial intelligence paradigms. The strong internal cohesion of this cluster signifies a transition toward representation learning, network reasoning, and intelligent automation, which have redefined the frontiers of computer and information systems in recent years.

The purple cluster structured around “social network”, “privacy”, and “community detection”, demonstrates KAIS’s responsiveness to societal and ethical dimensions of data science. It links algorithmic development to real-world contexts such as user interaction, data protection, and social behavior modeling, reflecting the journal’s expansion into trustworthy and human-centered AI.

Finally, the brown cluster—including “recommender systems”, “collaborative filtering”, and “matrix factorization”—and the green cluster, encompassing “semantic web” and “knowledge management”, represent emerging and niche research streams. These areas emphasize KAIS’s strength in supporting applied innovation and knowledge dissemination technologies, ensuring conceptual diversity across its portfolio.

Taken together, the thematic map illustrates a well-balanced intellectual ecosystem. Foundational areas such as data mining remain tightly connected to emerging frontiers in AI, knowledge engineering, and semantic computing. This structure highlights KAIS’s role as an integrative platform, fostering both continuity and innovation in the evolving landscape of intelligent information systems.

Figure 18 presents a word cloud depicting the most frequently used author keywords from the WoS database. This visualization, generated using Bibliometrix software, scales the size of each term in proportion to its frequency, providing an intuitive overview of the journal’s dominant research themes. The most prominent keywords— “data mining”, “classification”,



Fig. 18 Word cloud of the most frequent author keywords of KAIS

“clustering”, and “machine learning”—reflect the core methodological focus that has defined KAIS since its inception.

Beyond simple frequency counts, the word cloud underscores the conceptual coherence of the journal’s content: traditional topics in data analysis and pattern recognition coexist with emerging paradigms in artificial intelligence and knowledge-based computing. Collectively, these results highlight KAIS’s continuing emphasis on integrative, data-driven approaches to intelligent information systems and its central role in shaping methodological dialogue across disciplines.

To deepen the understanding of thematic evolution, this study incorporates a citation-driven topical analysis using SciVal, an analytical platform developed by Elsevier that provides insights into research performance, collaboration, and topic prominence based on Scopus data [84]. SciVal employs a range of bibliometric indicators, including the Prominence Percentile (PP) and the Field-Weighted Citation Impact (FWCI), to assess both the visibility and influence of research topics.

The Prominence Percentile [54] measures a topic’s momentum within the global research landscape, incorporating citation counts, usage metrics, and the average CiteScore of related publications. The FWCI [80] provides a normalized indicator of citation performance, comparing the citation rate of documents in KAIS with those of similar outputs indexed in Scopus, adjusted for publication year, document type, and subject area. An FWCI value of 1.0 indicates citation performance in line with the world average, while values above 1.0 signify greater-than-expected impact.

Table 15 summarizes the most frequently published topics in KAIS from 2013 to 2024, as identified through SciVal’s topical classification. This analysis complements the keyword-based findings, revealing how KAIS’s research themes evolve not only conceptually but also in terms of citation influence and global visibility, thus providing a dynamic perspective on the journal’s scholarly impact.

The top-ranked topic is “Association Rules, Frequent Itemsets, and Data Mining”, encompassing 32 publications, followed by “Time Series, Neural Network, and Data Mining” with

29 articles, and “Recommender Systems, Collaborative Filtering, and Graph Neural Network” with 27. Notably, the highest Field-Weighted Citation Impact (FWCI) corresponds to “Social Networking (Online), Sentiment Classification, and Data Mining” with an FWCI of 2.47, indicating that papers in this area receive more than twice the global average number of citations. Meanwhile, the topic “Graph Theory, Neural Network, and Graph Convolutional Network” exhibits the highest Prominence Percentile ($PP = 99.85$), reflecting exceptional global momentum and visibility.

Taken together, these findings demonstrate that KAIS’s intellectual core remains anchored in data mining and pattern discovery, while actively expanding toward graph-based learning, recommendation systems, and socially aware artificial intelligence. The coexistence of classical and emerging topics illustrates the journal’s dual identity as both a custodian of foundational methodologies and a catalyst for next-generation AI research, reaffirming its pivotal role in shaping the evolution of knowledge and information systems.

Ultimately, Table 16 presents the principal topic clusters in KAIS from 2014 to 2023, as identified through the Scopus database using the SciVal platform. This analysis employs the same indicators as Table 15—the Prominence Percentile (PP) and the Field-Weighted Citation Impact (FWCI)—to provide a multidimensional view of thematic performance.

The cluster “Data Mining, Graph Neural Networks, and Social Network Analysis” ranks first, comprising 139 publications, followed by “Sentiment Analysis, Natural Language Processing, and Machine Learning” with 115 papers, and “Data Mining, Artificial Intelligence, and Information Systems” in third position with 105 papers. In terms of citation performance, “Number, Multiple-Criteria Decision Analysis, and Fuzzy Logic” achieves the highest FWCI (2.08), indicating that its papers are cited more than twice the global average. Meanwhile, “Sentiment Analysis, Natural Language Processing, and Machine Learning” displays the highest Prominence Percentile ($PP = 99.67$), signifying strong global momentum and sustained scholarly attention.

Collectively, these results highlight the thematic convergence between traditional domains such as data mining and decision analysis and the rapidly expanding frontiers of machine learning, graph-based reasoning, and computational linguistics. This convergence demonstrates KAIS’s adaptive evolution—bridging symbolic and statistical paradigms—and confirms its position as a key integrative platform for advancing research in intelligent, interpretable, and interconnected information systems.

7 Conclusions

KAIS has completed 25 years of publications. To celebrate this milestone, a comprehensive bibliometric analysis was conducted for the period 1999–2024, examining publication and citation trends, intellectual influences, and evolving research themes. Drawing on data from the WoS Core Collection, the study identifies the most productive authors, institutions, and countries, as well as the most frequently cited documents and keywords that have shaped the journal’s intellectual trajectory.

From a theoretical standpoint, this research contributes to the understanding of how knowledge and information systems evolve as an interdisciplinary field, integrating principles of bibliometrics, data science, and artificial intelligence. It establishes an evidence-based framework that traces the intellectual structure and thematic progression of KAIS over twenty-five years, highlighting the journal’s role in consolidating methodological rigor and fostering cross-domain innovation. Methodologically, the study advances a replicable bibliometric

workflow that combines classical indicators with science mapping, co-citation, and network analyses using VOSviewer and Bibliometrix. This integrative approach enhances transparency, reproducibility, and comparative potential for future longitudinal evaluations of academic journals. Collectively, these contributions reinforce KAIS's position as both a subject of inquiry and a benchmark for studying the evolution of scientific knowledge in the era of data-driven research.

7.1 General findings

The results demonstrate KAIS's sustained growth and increasing global influence over the past quarter-century. Since its inception, the journal has published 2,385 articles, collectively accumulating 54,319 citations. The most influential contributions include the seminal works of Wu et al. [98] and Keogh & Ratanamahatana [49], with 3,427 and 1,176 citations respectively. Across the broader citation landscape, foundational papers by Agrawal, Demser, Blei, and Wu have emerged as cornerstones of KAIS's intellectual identity.

Philip S. Yu stands as the most prolific author, followed by Jian Pei and Jianzhong Li, reflecting the journal's dual emphasis on methodological innovation and practical application. At the institutional level, the Chinese Academy of Sciences and IBM lead in productivity and citation impact, while China and the United States dominate national contributions, complemented by strong participation from India, Australia, and Spain.

The VOSviewer analyses reveal a cohesive and interconnected intellectual landscape, highlighting three dominant bibliometric dimensions: co-citation, bibliographic coupling, and keyword co-occurrence. KAIS exhibits strong bibliographic ties to top-tier journals such as the *IEEE Transactions on Knowledge and Data Engineering*, *Expert Systems with Applications*, the *Journal of Machine Learning Research*, and *Information Sciences*. Within KAIS itself, self-referential co-citations indicate a mature research community with a strong internal knowledge base.

Keyword analyses identify enduring themes—data mining, classification, clustering, and machine learning—alongside emerging frontiers including deep learning, graph neural networks, natural language processing, and concept drift. Results from SciVal (2014–2023) confirm these patterns, showing that the journal's most active research clusters align with machine learning, artificial intelligence, social network analysis, and graph theory, reflecting KAIS's evolution from classical data mining to data-centric AI systems.

Finally, KAIS enjoys growing international recognition and citation impact across diverse domains such as computer science, engineering, telecommunications, mathematics, and operations research. Its consistent performance in WoS and Scopus rankings consolidates its position as a leading venue in the Information Systems and Artificial Intelligence categories.

7.2 Discussion

The findings provide several broader insights into the evolution and positioning of KAIS within the global research ecosystem. First, the journal's success lies in its ability to bridge traditional data mining methodologies with modern AI paradigms. The sustained prominence of topics such as classification and clustering underscores its foundational role in shaping the core algorithms of intelligent systems, while the recent rise of deep learning and graph neural networks signals its adaptability to contemporary computational challenges.

When compared with global trends in artificial intelligence and data science, the thematic evolution of KAIS shows both alignment and selectivity. While the global research landscape has rapidly shifted toward deep learning, generative models, and large-scale foundation architectures, KAIS has incorporated these paradigms more gradually, maintaining a strong emphasis on methodological evaluation, data-centric learning, and knowledge-oriented applications. This selective integration reflects the journal's traditional role as a bridge between theoretical advances and practical knowledge systems.

Second, the geographical distribution of contributions reflects the ongoing globalization of scientific inquiry. While early development was dominated by North American institutions, the more recent ascendancy of Chinese and Asia-Pacific research networks demonstrates a significant shift in knowledge production. This transformation underscores KAIS's role as a catalyst for international collaboration, connecting leading universities and research centres across continents.

Third, collaborative patterns observed through bibliographic coupling and institutional networks point to a maturing ecosystem in which multidisciplinary partnerships are essential. The journal's integration of knowledge engineering, data systems, and decision support demonstrates the convergence of previously distinct research traditions into unified frameworks for intelligent systems.

Finally, emerging themes related to interpretability, fairness, and trustworthiness reflect a new stage of methodological and ethical maturity in AI research. As intelligent systems increasingly influence decision-making, KAIS has positioned itself not only as a venue for technical advancement but also as a platform promoting responsible, explainable, and transparent knowledge discovery.

In essence, the findings underscore KAIS's dual mission: to preserve methodological rigor while fostering innovation in areas that redefine the boundaries of knowledge and information systems. Its trajectory mirrors that of the broader computational sciences—evolving from algorithmic precision toward data-driven intelligence and societal impact.

7.3 Practical implications

This study consolidates 25 years of scholarly contributions to KAIS, providing a structured understanding of the journal's intellectual development. By mapping its bibliometric trajectory, the paper offers researchers, educators, and editors a comprehensive overview of key authors, concepts, and collaborations that define the field. The identification of core research clusters and emerging themes can inform future editorial strategies, guiding the design of special issues, thematic calls, and collaborative initiatives in rapidly evolving areas such as explainable AI, graph-based learning, and knowledge graph reasoning.

Moreover, the methodological framework developed in this study offers a replicable model for bibliometric evaluation, facilitating comparative analyses across related journals and enabling longitudinal assessments of disciplinary growth within the broader landscape of information systems research.

7.4 Recommendations for future editorial and research directions

Based on the bibliometric trajectory observed over twenty-five years, several strategic directions emerge for the continued development of KAIS.

First, trustworthy and responsible AI appears as a unifying theme linking interpretability, fairness, privacy, and data governance. Future research could focus on evaluation frameworks, transparency protocols, and accountable AI systems in knowledge-intensive domains.

Second, data governance and reproducibility constitute an increasingly critical area. Studies addressing data quality, benchmarking standards, and lifecycle management of knowledge systems may reinforce KAIS's methodological foundations.

Third, generative and foundation models represent a major frontier in artificial intelligence. KAIS is well positioned to host research on generative models for structured knowledge, graph-based generation, and hybrid symbolic–neural architectures.

Fourth, the integration of graph learning, knowledge graphs, and decision intelligence systems suggests opportunities for cross-domain applications in policy, healthcare, and digital governance.

These directions align with both the journal's historical emphasis on methodological rigor and the emerging priorities of the global AI research landscape.

7.5 Limitations and future research

Despite offering a holistic overview, this study has certain limitations. First, the analysis relies primarily on the WoS Core Collection, which—although widely used in bibliometric research—may omit relevant outputs indexed in other databases. This limitation is particularly important in computer science, where high-impact contributions are often disseminated through conference proceedings. Compared with sources such as Scopus or DBLP Computer Science Bibliography, WoS provides more limited and uneven coverage of conference literature across subfields, which may affect the visibility of certain research streams and citation patterns.

Moreover, the predominance of English-language journals in WoS may introduce regional and linguistic biases. Contributions published in non-English outlets or regionally focused venues may therefore be underrepresented, potentially influencing country- and institution-level comparisons.

Second, WoS applies a full-counting approach by default, which can inflate the apparent contribution of multi-authored publications. To mitigate this effect, fractional counting was applied in the VOSviewer mapping analyses.

In addition, disciplinary differences in publication and citation behaviour complicate direct comparisons across subfields. Some findings—particularly those related to emerging research areas—may evolve rapidly, highlighting the need for periodic bibliometric updates, for example at the journal's 30th or 40th anniversaries.

Future research could extend the present analysis through cross-database comparisons (e.g., Scopus or Dimensions) and by incorporating alternative impact indicators such as altmetrics. Integrating full-text text mining, topic modelling, and longitudinal thematic analyses may further strengthen the identification of conceptual shifts and emerging research trajectories, particularly in areas aligned with KAIS's editorial focus, such as trustworthy AI, data governance, and generative knowledge systems.

Acknowledgements We would like to thank the editor-in-chief of KAIS, Xindong Wu, and the anonymous reviewers, for valuable comments that have significantly improved the quality of the paper.

Author contributions L.B.P. wrote the main manuscript text and prepared Tables 2 and 8, 9, 10, 11 and 12. J.M.M. conceived the idea, supervised the research, and prepared Tables 1, 3, 4, 5, 6 and 7, and 13, 14, 15 and 16, and all the figures. A.M.G. supervised the research. All authors reviewed the manuscript.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Data availability The data was generated using the Web of Science Core Collection and Scopus databases.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Agrawal RS, Srikant R (1994) R. Fast algorithms for mining association rules. *Proc 20th Int Conf Very Large Data Bases (VLDB)*, 487–499.
2. Agrawal R, Imieliński T, Swami A (1993) Mining association rules between sets of items in large databases. *ACM SIGMOD Rec* 22(2):207–216. <https://doi.org/10.1145/170036.170072>
3. Alonso S, Cabrerizo FJ, Herrera-Viedma E, Herrera F (2009) H-index: a review focused on its variants, computation, and standarization for different scientific fields. *J Informetrics* 3(4):273–289. <https://doi.org/10.1016/j.joi.2009.04.001>
4. Aminikhanghahi S, Cook DJ (2017) A survey of methods for time series change point detection. *Knowl Inf Syst* 51:339–367. <https://doi.org/10.1007/s10115-016-0987-z>
5. Aria M, Cuccurullo C (2017) Bibliometrix: an R-tool for comprehensive science mapping analysis. *J Informetrics* 11(4):959–975. <https://doi.org/10.1016/j.joi.2017.08.007>
6. Arrow KJ, Blinder AS, Solow RM, Stiglitz JE (2011) 100 years of the *American Economic Review*: the top 20 articles. *Am Econ Rev* 101(1):1–8. <https://doi.org/10.1257/aer.101.1.1>
7. Bar-Ilan J (2008) Informetrics at the beginning of the 21st century: a review. *J Informetrics* 2(1):1–52. <https://doi.org/10.1016/j.joi.2007.11.001>
8. Barcellos-Paula L, Merigó JM, Gil-Lafuente AM (2024) 100 volumes of mathematical methods of operations research: a bibliometric overview. *Math Methods Oper Res* 100(3):753–796. <https://doi.org/10.1007/s00186-024-00883-y>
9. Bensman SJ (2007) Garfield and the impact factor. *Annu Rev Inf Sci Technol* 41(1):93–155. <https://doi.org/10.1002/aris.2007.1440410110>
10. Bergstrom CT, West JD, Wiseman MA (2008) The Eigenfactor™ metrics. *J Neurosci* 28(45):11433–11434. <https://doi.org/10.1523/JNEUROSCI.0003-08.2008>
11. Bhukya R, Verma S, Kumar S, Donthu N (2021) Forty years of *European Management Journal*: a bibliometric overview. *Eur Manag J* 39(6):739–752. <https://doi.org/10.1016/j.emj.2021.05.004>
12. Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet allocation. *J Mach Learn Res* 3:993–1022
13. Bolón-Canedo V, Sánchez-Marroño N, Alonso-Betanzos A (2013) A review of feature selection methods on synthetic data. *Knowl Inf Syst* 34:483–519. <https://doi.org/10.1007/s10115-012-0487-8>
14. Broadus RN (1987) Toward a definition of “bibliometrics”. *Scientometrics* 12:373–379. <https://doi.org/10.1007/BF02016680>
15. Callon M, Courtial JP, Turner WA, Bauin S (1983) From translations to problematic networks: an introduction to co-word analysis. *Soc Sci Inf* 22(2):191–235. <https://doi.org/10.1177/053901883022002003>
16. Calma A, Ho W, Shao LS, Li HS (2021) *Operations research*: topics, impact, and trends from 1952–2019. *Oper Res* 69(5):1487–1508. <https://doi.org/10.1287/opre.2021.2139>
17. Cancino CA, Merigó JM, Coronado F, Dessouky Y, Dessouky M (2017) Forty years of *computers & industrial engineering*: a bibliometric analysis. *Comput Ind Eng* 113:614–629. <https://doi.org/10.1016/j.cie.2017.08.033>

18. Clarivate (2024). Journal citation reports: reference guide. J Cit Rep, 11. <https://clarivate.com/products/web-of-science/>
19. Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F (2011) An approach for detecting, quantifying, and visualizing the evolution of a research field: a practical application to the fuzzy sets theory field. J Informetr 5(1):146–166. <https://doi.org/10.1016/j.joi.2010.10.002>
20. Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F (2012) SciMAT: a new science mapping analysis software tool. J American Soc Inform Sci Tech 63(8):1609–1630. <https://doi.org/10.1002/asi.22688>
21. Cobo M, Martínez M, Gutiérrez-Salcedo M, Fujita H, Herrera-Viedma E (2015) 25 years at knowledge-based systems: a bibliometric analysis. Knowl Based Syst 80:3–13. <https://doi.org/10.1016/j.knosys.2014.12.035>
22. Coello CAA (1999) Comprehensive survey of evolutionary-based multiobjective optimization techniques. Knowl Inf Syst 1:269–308. <https://doi.org/10.1007/BF03325101>
23. Cooley R, Mobasher B, Srivastava J (1999) Data preparation for mining World Wide Web browsing patterns. Knowl Inf Syst 1:5–32. <https://doi.org/10.1007/BF03325089>
24. Demšar J (2006) Statistical comparisons of classifiers over multiple data sets. J Mach Learn Res 7:1–30
25. Ding Y, Rousseau R, Wolfram D (2014) Measuring Scholarly Impact: methods and practice. In Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Representing Scientific Knowledge* (1st ed., Issue 1982). Springer Cham. <https://doi.org/10.1007/978-3-319-10377-8>
26. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM (2020) Journal of business research: a 45-year retrospective analysis. J Business Research 109:1–15. <https://doi.org/10.1016/j.jbusres.2020.02.042>
27. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM (2021) How to conduct a bibliometric analysis: an overview and guidelines. J Bus Res 133:285–296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
28. Duong H, Truong T, Le B, Fournier-Viger P (2024) CG-FHAUL: an efficient algorithm for simultaneously mining succinct pattern sets of frequent high average utility itemsets. Knowl Inf Syst 66:5239–5280. <https://doi.org/10.1007/s10115-024-02121-7>
29. Ezugwu AE, Shukla AK, Nath R, Akinyelu AA, Agushaka JO, Chiroma H, Muhuri PK (2021) Meta-heuristics: a comprehensive overview and classification along with bibliometric analysis. Artif Intell Rev 54(6):4237–4316. <https://doi.org/10.1007/s10462-020-09952-0>
30. Figuerola-Wischke A, Merigó JM, Gil-Lafuente AM, Kydland FE, Amiguet L (2024) The Scandinavian Journal of Economics at 125: a bibliometric overview. Scand J Econ 126(4):643–697. <https://doi.org/10.1111/sjoe.12582>
31. García-Merino MT, Pereira-do-Carmo ML, Santos-Álvarez MV (2006) 25 years of Technovation: characterisation and evolution of the journal. Technovation 26(12):1303–1316. <https://doi.org/10.1016/j.technovation.2005.11.005>
32. García-Orozco D, Alfaro-García VG, Merigó JM, Moreno ICE, Monge RG (2022) An overview of the most influential journals in fuzzy systems research. Expert Syst Appl 200:117090. <https://doi.org/10.1016/j.eswa.2022.117090>
33. Gardner WL, Lowe KB, Meuser JD, Noghani F, Gullifor DP, Coglisier CC (2020) The leadership trilogy: a review of the third decade of The Leadership Quarterly. Leadersh Q 31(1):101379. <https://doi.org/10.1016/j.leaqua.2019.101379>
34. Garfield E (1955) Citation indexes for science: a new dimension in documentation through association of ideas. Science 122(3159):108–111. <https://doi.org/10.1126/science.122.3159.108>
35. Gedik B, Wu KL, Yu PS, Liu L (2007) CPU load shedding for binary stream joins. Knowl Inf Syst 13:271–303. <https://doi.org/10.1007/s10115-006-0044-4>
36. Glänzel W, Moed HF, Schubert A, Thelwall M (2019) Bibliometrics as a research field: a course on theory and application of bibliometric indicators. Scientometrics 121(3):1819–1838. <https://doi.org/10.1007/s11192-019-03207-4>
37. Gorsky M, Arnold-Forster A (2023) The Lancet at 200: medicine, media, and modernity. Lancet 401(10375):1639–1640. [https://doi.org/10.1016/S0140-6736\(23\)00710-8](https://doi.org/10.1016/S0140-6736(23)00710-8)
38. Guo YM, Huang ZL, Guo J, Li H, Liu MY (2021) A bibliometric analysis and visualization of blockchain. Future Gener Comput Syst 116:316–332. <https://doi.org/10.1016/j.future.2020.10.023>
39. Hart OD, Mizon GE (1983) Editorial note: the review of economic studies—past, present and future. Rev Econ Stud 50(4):555–557. <https://doi.org/10.2307/2297762>
40. Heck JL, Cooley PL, Hubbard CM (1986) An analysis of published research in the *Journal of Finance*: 1946–1985. J Finance 41(5):1129–1140. <https://doi.org/10.1111/j.1540-6261.1986.tb04526.x>
41. Hirsch JE (2005) An index to quantify an individual's scientific research output. Proc Natl Acad Sci USA 102(46):16569–16572. <https://doi.org/10.1073/pnas.0507655102>
42. Hu X, Chu L, Pei J, Liu W, Bian J (2021) Model complexity of deep learning: a survey. Knowl Inf Syst 63:2585–2619. <https://doi.org/10.1007/s10115-021-01605-0>

43. Huang C, Lin M, Xu Z (2020) Pythagorean fuzzy MULTIMOORA method based on distance measure and score function: its application in multicriteria decision making process. *Knowl Inf Syst* 62:4373–4406. <https://doi.org/10.1007/s10115-020-01491-y>
44. Inkpen AC, Beamish PW (1994) An analysis of 25 years of research in the journal of international-business studies. *J Int Bus Stud* 25(4):703–713. <https://doi.org/10.1057/palgrave.jibs.8490220>
45. Jain AK, Murty MN, Flynn PJ (1999) Data clustering: a review. *ACM Comput Surv* 31(3):264–323. <https://doi.org/10.1145/331499.331504>
46. Jiang M, Cui P, Beutel A, Faloutsos C, Yang S (2016) Inferring lockstep behavior from connectivity pattern in large graphs. *Knowl Inf Syst* 48:399–428. <https://doi.org/10.1007/s10115-015-0883-y>
47. Kamiran F, Calders T (2012) Data preprocessing techniques for classification without discrimination. *Knowl Inf Syst* 33:1–33. <https://doi.org/10.1007/s10115-011-0463-8>
48. Keogh E, Lin J (2005) Clustering of time-series subsequences is meaningless: implications for previous and future research. *Knowl Inf Syst* 8:154–177. <https://doi.org/10.1007/s10115-004-0172-7>
49. Keogh E, Ratanamahatana C (2005) Exact indexing of dynamic time warping. *Knowl Inf Syst* 7:358–386. <https://doi.org/10.1007/s10115-004-0154-9>
50. Keogh E, Chakrabarti K, Pazzani M, Mehrotra S (2001) Dimensionality reduction for fast similarity search in large time series databases. *Knowl Inf Syst* 3:263–286. <https://doi.org/10.1007/PL00011669>
51. Kessler MM (1963) Bibliographic coupling between scientific papers. *Am Doc*. <https://doi.org/10.1002/asi.5090140103>
52. Kirchler E, Hölzl E (2006) Twenty-five years of the *Journal of Economic Psychology* (1981–2005): A report on the development of an interdisciplinary field of research. *J Econ Psychol* 27(6):793–804. <https://doi.org/10.1016/j.joep.2006.09.005>
53. Kirişçi M (2023) New cosine similarity and distance measures for Fermatean fuzzy sets and TOPSIS approach. *Knowl Inf Syst* 65:855–868. <https://doi.org/10.1007/s10115-022-01776-4>
54. Klavans R, Boyack KW (2017) Research portfolio analysis and topic prominence. *J Informetr* 11(4):1158–1174. <https://doi.org/10.1016/j.joi.2017.10.002>
55. Koca M, Çiftçi S (2025) A comprehensive bibliometric analysis of Big Data and Cyber Security: intellectual structure, trends, and global collaborations. *Knowl Inf Syst*. <https://doi.org/10.1007/s10115-025-02531-1>
56. Kokol P, Vošner HB, Završnik J (2020) Application of bibliometrics in medicine: a historical bibliometrics analysis. *Health Info Libr J*. <https://doi.org/10.1111/hir.12295>
57. Kube R, von Gradowski R, Jakob M (2018) The evolution of environmental economics: a bibliometric analysis of *Journal of Environmental Economics and Management* (1974–2017). *J Environ Econ Manag* 90:1–14. <https://doi.org/10.1016/j.jeem.2018.05.003>
58. Kumar S, Lim WM, Paul J (2022) *Journal of International Management* at 25: a bibliometric analysis and future research directions. *J Int Manag* 28(4):100954. <https://doi.org/10.1016/j.intman.2022.100954>
59. Kumar S, Pandey N, Mukherjee D (2022) Cross cultural and strategic management: a retrospective overview using bibliometric analysis. *Cross Cult Strateg Manag* 29(1):171–194. <https://doi.org/10.1108/ccsm-08-2021-0147>
60. Laengle S, Lobos V, Merigó JM, Herrera-Viedma E, Cobo MJ, De Baets B (2021) Forty years of *Fuzzy Sets and Systems*: a bibliometric analysis. *Fuzzy Sets Syst* 402:155–183. <https://doi.org/10.1016/j.fss.2020.03.012>
61. Lee, L.K., Webb, G.I., Schmidt, D.F., & Piatkowski (2024). Computing marginal and conditional divergences between decomposable models with applications in quantum computing and earth observation. *Knowl Inf Syst*, 66, 7527–7556. <https://doi.org/10.1007/s10115-024-02191-7>
62. Li C, Jiang L, Li H, Wu J, Zhang P (2017) Toward value difference metric with attribute weighting. *Knowl Inf Syst* 50:795–825. <https://doi.org/10.1007/s10115-016-0960-x>
63. Li X, Xiong H, Li X, Wu X, Zhang X, Liu J, Bian J, Dou D (2022) Interpretable deep learning: interpretation, interpretability, trustworthiness, and beyond. *Knowl Inf Syst* 64:3197–3234. <https://doi.org/10.1007/s10115-022-01756-8>
64. Li Y, Xu ZS, Wang XX, Wang XZ (2020) A bibliometric analysis on deep learning during 2007–2019. *Int J Mach Learn Cybern* 11(12):2807–2826. <https://doi.org/10.1007/s13042-020-01152-0>
65. Liu J, Huang J, Zhou Y, Li X, Ji S, Xiong H, Dou D (2022) From distributed machine learning to federated learning: a survey. *Knowl Inf Syst* 64:885–917. <https://doi.org/10.1007/s10115-022-01664-x>
66. López-Robles JR, Cobo MJ, Gutiérrez-Salcedo M, Martínez-Sánchez MA, Gamboa-Rosales NK, Herrera-Viedma E (2021) 30th anniversary of Applied Intelligence: a combination of bibliometrics and thematic analysis using SciMAT. *Appl Intell* 51(9):6547–6568. <https://doi.org/10.1007/s10489-021-02584-z>

67. Martorell Cunill O, Socías Salvá A, Otero Gonzalez L, Mulet-Forteza C (2019) Thirty-fifth anniversary of the International Journal of Hospitality Management: a bibliometric overview. *Int J Hosp Manag* 78:89–101. <https://doi.org/10.1016/j.ijhm.2018.10.013>
68. Merigó JM, Blanco-Mesa F, Gil-Lafuente AM, Yager RR (2017) Thirty years of the *International Journal of Intelligent Systems*: a bibliometric review. *Int J Intell Syst* 32(5):526–554. <https://doi.org/10.1002/int.21859>
69. Merigó JM, Pedrycz W, Weber R, de la Sotta C (2018) Fifty years of *Information Sciences*: a bibliometric overview. *Inf Sci* 432:245–268. <https://doi.org/10.1016/j.ins.2017.11.054>
70. Merigó JM, Yang JB (2017) A bibliometric analysis of operations research and management science. *Omega-Int J Manage Sci* 73:37–48. <https://doi.org/10.1016/j.omega.2016.12.004>
71. Mikolov T, Corrado G, Chen K, Dean J (2013) word2vec-v1. *Proc Int Conf Learning Representations (ICLR 2013)*, 1–12.
72. Monastersky R, Van Noorden R (2019) *Nature* at 150: the past, present and future of the world's leading science journal. *Nature* 575(7782):574–577. <https://doi.org/10.1038/d41586-019-03584-7>
73. Mukhtiar HZ, Mustafa G, Afzal MT (2025) Enhancing researcher evaluation in computer science: a novel index for impact assessment. *Knowl Inf Syst*. <https://doi.org/10.1007/s10115-025-02522-2>
74. Mustafa G, Afzal MT, Khalid YN (2025) Knowledge discovery through interpretable decision rules: a framework for ranking researchers from bibliometric data. *Knowl Inf Syst*. <https://doi.org/10.1007/s10115-025-02597-x>
75. Narayanan K, Manimaran A (2024) Recent developments in geographic information systems across different application domains: a review. *Knowl Inf Syst* 66(3):1523–1547. <https://doi.org/10.1007/s10115-024-01956-9>
76. Paul J, Lim WM, O'Cass A, Hao AW, Bresciani S (2021) Scientific procedures and rationales for systematic literature reviews (SPAR-4-SLR). *Int J Consum Stud*. <https://doi.org/10.1111/ijcs.12695>
77. Paz A, Merigó J, Powell P, Ramaprasad A, Syn T (2019) Twenty-five years of the information systems journal: a bibliometric and ontological overview. *Inf Syst J* 30:431–457. <https://doi.org/10.1111/isj.12260>
78. Pons P, Latapy M (2006) Computing communities in large networks using random walks. *J Graph Algorithms Appl* 10(2):191–218. <https://doi.org/10.7155/jgaa.00124>
79. Pritchard A (1969) Statistical bibliography or bibliometrics? *J Doc* 25(4):348–349
80. Purkayastha A, Palmaro E, Falk-Krzesinski HJ, Baas J (2019) Comparison of two article-level, field-independent citation metrics: field-weighted citation impact (FWCI) and relative citation ratio (RCR). *J Informetr* 13(2):635–642. <https://doi.org/10.1016/j.joi.2019.03.012>
81. Rousseau R (2014) Forgotten founder of bibliometrics. *Nature* 510:218. <https://doi.org/10.1038/510218e>
82. Saha B, Gupta S, Phung D, Venkatesh S (2017) Effective sparse imputation of patient conditions in electronic medical records for emergency risk predictions. *Knowl Inf Syst* 53:179–206. <https://doi.org/10.1007/s10115-017-1038-0>
83. Schwert GW (1993) The journal of financial economics: a retrospective evaluation (1974–91). *J Financ Econ* 33:369–424. [https://doi.org/10.1016/0304-405x\(93\)90012-z](https://doi.org/10.1016/0304-405x(93)90012-z)
84. SciVal. (2024). *Quick reference guide*.
85. Scopus. (2024). *Scopus Database*.
86. Shukla AK, Janmajaya M, Abraham A, Muhuri PK (2019) Engineering applications of artificial intelligence: a bibliometric analysis of 30 years (1988–2018). *Eng Appl Artif Intell* 85:517–532. <https://doi.org/10.1016/j.engappai.2019.06.010>
87. Small H (1973) Co-citation in the scientific literature: a new measure of the relationship between two documents. *J Am Soc Inf Sci*. <https://doi.org/10.1002/asi.4630240406>
88. Štrumbelj E, Kononenko I (2014) Explaining prediction models and individual predictions with feature contributions. *Knowl Inf Syst* 41:647–665. <https://doi.org/10.1007/s10115-013-0679-x>
89. Tang G, Pei J, Luk WS (2014) Email mining: tasks, common techniques, and tools. *Knowl Inf Syst* 41:1–31. <https://doi.org/10.1007/s10115-013-0658-2>
90. Thompson DF, Walker CK (2015) A descriptive and historical review of bibliometrics with applications to medical sciences. *Pharmacotherapy* 35(6):551–559. <https://doi.org/10.1002/phar.1586>
91. Tukey JW (1977) *Exploratory data analysis* (1st ed., Vol. 1). Addison-Wesley.
92. Tzika-Kostopoulou D, Nathanail E, Kokkinos K (2024) Big data in transportation: a systematic literature analysis and topic classification. *Knowl Inf Syst* 66:5021–5046. <https://doi.org/10.1007/s10115-024-02112-8>
93. Van Eck NJ, Waltman L (2007) Bibliometric mapping of the computational intelligence field. *Int J Uncertain Fuzziness Knowl Based Syst* 15(5):625–645. <https://doi.org/10.1142/S0218488507004911>
94. Van Eck NJ, 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>

95. Van Noorden R, Maher B, Nuzzo R (2014) The top 100 papers. *Nature*. <https://doi.org/10.1038/514550a>
96. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L & Polosukhin I (2017) Attention is all you need. *Proc 31st Int Conf Neural Information Processing Systems*, 6000–6010.
97. Wang XX, Chang YR, Xu ZS, Wang ZD, Kadirkamanathan V (2021) 50 years of International Journal of Systems Science: a review of the past and trends for the future. *Int J Syst Sci* 52(8):1515–1538. <https://doi.org/10.1080/00207721.2020.1862937>
98. Wu X, Kumar V, Quinlan J, Ghosh J, Yang Q, Motoda H, McLachlan G, Ng A, Liu B, Yu P, Zhou Z-H, Steinbach M, Hand D, Steinberg D (2008) Top 10 algorithms in data mining. *Knowl Inf Syst* 14:1–37. <https://doi.org/10.1007/s10115-007-0114-2>
99. Xu ZS, Ge ZJ, Wang XX, Kou G (2023) A look at the focus on big data for information technology and decision making during 1994 to 2020. *Int J Inf Technol Decis Mak* 22(1):7–35. <https://doi.org/10.1142/S0219622022500298>
100. Xu ZS, Yu DJ, Wang XZ (2019) A bibliometric overview of international journal of machine learning and cybernetics between 2010 and 2017. *Int J Mach Learn Cybern* 10(9):2375–2387. <https://doi.org/10.1007/s13042-018-0875-9>
101. Yang J, Leskovec J (2015) Defining and evaluating network communities based on ground-truth. *Knowl Inf Syst* 42: 181–213. <https://doi.org/10.1007/s10115-013-0693-z>
102. Yazdanjue N, Yazdanjoui H, Gharoun H, Khorshidi M, Rakhshaninejad M, Amiri B, Gandomi A (2025) A comprehensive bibliometric analysis on social network anonymization: current approaches and future directions. *Knowl Inf Syst* 67:29–108. <https://doi.org/10.1007/s10115-024-02289-y>
103. Yu DJ, Xu ZS, Kao YS, Lin CT (2018) The structure and citation landscape of IEEE transactions on fuzzy systems (1994–2015). *IEEE Trans Fuzzy Syst* 26(2):430–442. <https://doi.org/10.1109/tfuzz.2017.2672732>
104. Yu DJ, Xu ZS, Martínez L (2020) Visualizing the intellectual structure of the fuzzy linguistic knowledge domain: a bibliometric analysis. *Int J Fuzzy Syst* 22(8):2397–2413. <https://doi.org/10.1007/s40815-020-00959-x>
105. Yu DJ, Xu ZS, Wang WR (2019) A bibliometric analysis of fuzzy optimization and decision making (2002–2017). *Fuzzy Optim Decis Mak* 18(3):371–397. <https://doi.org/10.1007/s10700-018-9301-8>
106. Zheng K, Wang H, Qi Z, Li J, Gao H (2017) A survey of query result diversification. *Knowl Inf Syst* 51:1–36. <https://doi.org/10.1007/s10115-016-0990-4>
107. Zurita G, Merigó JM, Lobos-Ossandón V, Mulet-Forteza C (2021) Leading countries in computer science: a bibliometric overview. *J Intell Fuzzy Syst* 40(2):1957–1970. <https://doi.org/10.3233/jifs-189199>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Luciano Barcellos-Paula is an Associate Professor at the CENTRUM PUCP from the Pontifical Catholic University of Peru. He holds a PhD in Business Administration, a PostDoc, and a Master's Degree in Business, Finance and Insurance from the University of Barcelona (Spain), a MBA in Business Administration from the Polytechnic University of Catalonia, a MBA & Leisure Management, Business Administration from the EAE Business School (Spain), and a Bachelor's degree in Economic Sciences from the Federal University of Fluminense (Brazil). He is Corresponding Academic Member elected by Brazil of the Spanish Royal Academy of Economic and Financial Sciences. Director of the Observatory of Innovation and Sustainability at CENTRUM PUCP and Chair of the Center for Sustainability & Social Innovation at CENTRUM PUCP. He has published more than 100 papers in journals, books and conference proceedings including journals such as International Journal of Intelligent Systems, Journal of Cleaner Production, Journal of Strategic Marketing, Corporate Social Responsibility and Environmental Management, Mathematical Methods of Operations Research, Kybernetes, Cybernetics & Systems, International Journal Uncertainty, Fuzziness and Knowledge-Based Systems, Sustainable Production and Consumption and Journal of Public Affairs. He is interested in sustainability and stakeholder engagement, corporate social responsibility, corporate governance, fuzzy logic, bibliometrics and decision-making. Webpage: https://scholar.google.com/citations?hl=es&user=mJ06Y_YAAAAJ&view_op=list_works.

José M. Merigó is a Full Professor at the School of Computer Science, Faculty of Engineering and Information Technology, University of Technology Sydney (Australia). Before, he was a Full Professor at the University of Chile, Senior Research Fellow at the University of Manchester (UK) and Assistant Professor at the University of Barcelona (Spain). He holds a Bachelor, Master and a PhD degree in Business Administration from the University of Barcelona, and a BSc and MSc degree from Lund University (Sweden). He

has published more than 500 articles in journals, books, and conference proceedings. He is or has been on the editorial board of many journals including the Journal of Management & Organization, Computers & Industrial Engineering, Journal of Business Research, PLOS One, Humanities and Social Sciences Communications, Kybernetes and International Journal of Fuzzy Systems. Recently, Clarivate Analytics (previously Thomson & Reuters) has distinguished him as a Highly Cited Researcher in Computer Science (2015–2020) and Cross-Field (2021–2024). He has also appeared in many other international rankings including the Stanford/Elsevier Top 2% Scientists List, Research.com, ScholarGPS, and AD Scientific Index. He is currently interested in decision making, aggregation systems, computational intelligence, business analytics, bibliometrics, and applications in business, economics, and social sciences. Webpage: <http://scholar.google.cl/citations?user=rprb9KwAAAAJ&hl=es>.

Ana M. Gil-Lafuente is a Full Professor at the Department of Business Administration at the University of Barcelona (Spain). She has a MSc and a PhD Degree in Business Administration from the University of Barcelona. She has published more than 500 papers in journals, books and conference proceedings including journals such as Information Sciences, International Journal of Intelligent Systems, Applied Soft Computing, Scandinavian Journal of Economics, International Journal of Approximate Reasoning, Land Economics, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems and Expert Systems with Applications. She has published more than 20 books most of them with Springer. She is co-editor-in-chief of 8 journals published by the Association for Modeling and Simulation in Enterprises (AMSE). She is on the editorial board of several journals and scientific committees. She has also served as a reviewer for many journals. She is currently interested in decision making, bibliometrics, aggregation operators, fuzzy systems, uncertainty and finance. Webpage: <https://scholar.google.com/citations?user=7r7JP2gAAAAJ&hl=es&oi=ao>.