Received: 18 May 2026 | Revised: 24 June 2026 | Accepted: 15 July 2026 | Published online: 30 July 2026

RESEARCH ARTICLE

Artificial Intelligence and Applications
2026, Vol. 00(00) 1–16
DOI: 10.47852/bonviewAIA620210405

BON VIEW PUBLISHING

Generative AI in Management Research After ChatGPT: A Bibliometric Map of Themes, Influence, and Governance

Irem Tanyıldızı Baydili[1], Burak Tasci[2,*], Sengul Dogan[3] and Turker Tuncer[3]

[1] *Faculty of Economics and Administrative Sciences, Firat University, Turkey*

[2] *Vocational School of Technical Sciences, Firat University, Turkey*

[3] *Department of Digital Forensics Engineering, Firat University, Turkey*

Abstract: Generative artificial intelligence (GenAI), particularly since the introduction of ChatGPT, has not only led to a quantitative increase in publications in management and business research but also created a qualitative transformation in research themes and collaborative structures. However, the current literature is insufficient to comprehensively present the epistemic and thematic consequences of this transformation in the field of management. This study aims to fill this gap by examining 316 articles indexed in the Social Sciences Citation Index database within the Web of Science Core Collection and published between 2023 and 2025 using bibliometric analysis. The analyses show that normative and governance-oriented themes such as ethics, artificial intelligence (AI) literacy, and agentic AI have become increasingly visible in the management literature in the post-ChatGPT era; rather than displacing performance- and productivity-centered approaches, these normative themes emerged alongside them within an increasingly diversified keyword and thematic structure. The findings reveal that GenAI research in management has evolved from being merely a technological innovation to a research axis that reshapes discussions on organizational responsibility, human–machine interaction, and governance. In this respect, the study goes beyond mapping the current state of the literature and offers theoretical and methodological implications on how GenAI should be positioned in management research.

Keywords: generative artificial intelligence, ChatGPT, bibliometric analysis, management and organization

1. Introduction

Research on management and organizations has focused on digital technologies since their introduction because scientists want to understand their impact on work productivity, organizational decision-making speed, and market competitiveness. Organizations gained a new method to understand digital resources through generative artificial intelligence (GenAI) because they could develop innovative digital instruments. The operation of large language models (LLMs) including ChatGPT differs from previous digital systems because they now serve as independent cognitive entities that help users generate knowledge, make decisions, and understand information. The changing business environment requires organizations to review their core management principles, which include decision authority, employee responsibility, ethical conduct, organizational leadership, and human–technology teamwork evolution.

The computer science subfield of artificial intelligence (AI) works to create systems that perform human intellectual tasks including learning, reasoning, problem-solving, perception, and natural language processing (NLP) [1]. The Dartmouth Workshop established AI as a formal field in 1956, yet scientists needed to wait multiple years to implement their discoveries into practical applications. The development of deep learning systems together with improved computing power and access to extensive datasets has created new possibilities for generative GenAI systems, which generate original content in different forms including text, images, code, audio, and video [2]. The creation of these systems requires extensive LLMs that generate statistical results that duplicate training data patterns and learn to function within particular situations [1].

The fast-changing environment contains three systems that demonstrate both the wide range of GenAI technologies and their quick spread throughout the market: ChatGPT [3, 4] operates through GPT-3 and GPT-4 technology, while Google uses LaMDA [5] and PaLM 2 [6] models under the Gemini framework, and DALL·E [7, 8] and Stable Diffusion operate as generative visual platforms [1]. The worldwide adoption of ChatGPT became possible when the system went public in November 2022, which demonstrated that GenAI functions as a strategic general-purpose technology that will transform economic

*Corresponding author: Burak Tasci, Vocational School of Technical Sciences, Firat University, Turkey. Email: btasci@firat.edu.tr

systems, organizational structures, and managerial practices. The fast adoption of this tool by millions of users demonstrates how knowledge work operations have evolved through new organizational systems and performance assessment methods.

GenAI systems operated by managers help humans perform their work better when they need to handle tasks that require language processing such as document summarization, presentation content creation, customer feedback evaluation, and decision-making system optimization [9]. The systems perform automated tasks that allow managers and staff members to redirect their mental abilities toward complex analytical work, creative thinking, and strategic planning. GenAI functions as a tool that organizations can use to create enduring market advantages through its various business applications including customer experience optimization, personalized marketing, advanced business intelligence, and data-driven human-resource management [10, 11]. The post-COVID-19 era has sped up this process because managers and organizations now understand digital technology better, which makes it easier for them to adopt and test AI systems [12].

Organizations that adopt GenAI technology face multiple structural weaknesses that produce both ethical challenges and administrative problems for their governance systems. Managerial issues have become essential because of scalability problems and workflow integration requirements, workforce changes and business model adjustments, algorithmic errors and fake information spread, privacy issues and security threats, and regulatory requirements [13, 14]. The identified risks show that GenAI needs more than technological solutions because it operates as an organizational decision system that needs assessment through accountability systems and governance frameworks.

Reflecting these developments, management scholarship has undergone a marked epistemic shift since the public release of ChatGPT. The number of GenAI-related publications has experienced a substantial increase since late 2022, while scientists now study multiple new research domains. Research studies three primary subjects, which focus on organizational knowledge generation and systems for authorship and peer review processes and human–AI system interactions. Research now focuses on creating responsible AI governance systems, which makes the reflexive direction more visible because organizations must train their employees about AI, create accountability systems for AI systems, study human interactions with AI systems, and define the criteria for AI system autonomy [14–16].

A critical gap remains in explaining how GenAI has reconfigured management scholarship after the public release of ChatGPT [15, 17–19]. Existing studies often treat AI as a broad and undifferentiated technological domain or focus on isolated applications, leaving unclear how post-ChatGPT research has reorganized its intellectual structure, influential actors, collaboration networks, and thematic priorities. This study addresses that gap by isolating Social Sciences Citation Index (SSCI)-indexed GenAI research in management, combining bibliometric mapping with statistical validation and robustness checks, and linking emerging themes such as agentic AI, ethics, AI literacy, and governance to managerial questions of accountability, autonomy, and organizational control.

Compared with existing AI reviews and bibliometric studies, this study offers three distinct contributions. Theoretically, it separates GenAI from broader AI research and positions it as a management-specific governance and agency issue after ChatGPT. Methodologically, it combines science mapping with statistical validation, robustness checks, and null-model benchmarking rather than relying only on descriptive bibliometric visualizations. Empirically, it shows that ethics, AI literacy, governance, and agentic AI have emerged alongside, rather than replaced, productivity-oriented themes in post-ChatGPT management scholarship. This study aims to examine the structural characteristics and research trends of the GenAI literature in the field of management using bibliometric analysis of Web of Science Core Collection data. Specifically, it aims to map the increasing academic interest in GenAI in management research in the post-ChatGPT era, revealing key trends, areas of focus, and research gaps in the literature. Importantly, this study does not examine the direct effects of GenAI on employees or organizational outcomes. Instead, it analyzes how post-ChatGPT management scholarship has framed GenAI and whether this literature remains productivity-oriented or increasingly incorporates ethics, governance, accountability, and AI literacy. Accordingly, the study seeks to answer the following research questions:

1) What is the distribution of publications by year?
2) Who are the most cited authors?
3) Which authors are most related according to their co-authorship network?
4) What is the structure of collaboration among authors?
5) What are the citation relationships between countries?
6) What are the most frequently used keywords?
7) What research themes have emerged over time?

2. Method

This study utilized bibliometric analysis, a quantitative research method. This bibliometric study analyzed 316 SSCI-indexed management and business articles published in the Web of Science Core Collection between 2023 and 2025, using VOSviewer(version 1.6.20)-based science mapping and a Python-based validation pipeline to examine keyword, author, country, citation, and collaboration structures. Bibliometric analysis is an effective analytical approach that aims to reveal the knowledge base, intellectual structure, and evolutionary dynamics over time of a particular research area by systematically, objectively, and reproducibly examining large volumes of often unstructured scientific data [18]. This method also makes it possible to reveal research trends, thematic concentrations, and inter-actor relationships in the literature based on the science-mapping approach [20].

Bibliometric analysis is an approach considered within the scope of informetric content-based analyses and provides a methodological framework for integrative literature reviews. Such analyses allow for the identification of prominent themes in the literature, the analysis of the intensity of relationships between themes, and the systematic presentation of conceptual structures. The main informetric analysis techniques commonly used in the literature include citation analysis, co-citation analysis, co-authorship analysis, bibliographic matching analysis, and keyword co-occurrence analysis. Each of these methods offers unique insights into different dimensions of the research area and possesses different strengths and limitations in terms of their analytical scope and the types of outputs they provide [19]. Methodologically sound bibliometric studies make significant contributions to the systematic identification of knowledge gaps in a research area and to creating a guiding foundation for future research [18].

In the bibliometric analysis of scientific literature, the use of Scopus and Web of Science (WoS) databases is recommended due to reasons such as the breadth of data coverage, record accuracy, and reduction of duplication. These databases enable the acquisition of reliable and high-quality datasets thanks to the stricter and more systematic criteria they apply in publication acceptance processes [19]. In this study, data were obtained from the Web of Science Core Collection database. Web of Science Core Collection was selected because it provides highly standardized bibliographic metadata, rigorous indexing procedures, and reliable citation-link structures, which are essential for reproducible bibliometric analysis [21, 22]. In line with bibliometric guidelines proposed by Donthu et al. [18], the use of a curated database helps reduce duplication, metadata inconsistency, and cross-platform indexing heterogeneity, thereby strengthening the internal validity of science-mapping analyses. Although Scopus provides broader coverage in some domains, the SSCI restriction in Web of Science was considered appropriate for identifying high-quality social science research directly related to management, business, business finance, and economics. References cited for theoretical background may include works outside the analyzed corpus; however, the bibliometric dataset itself was restricted to SSCI-indexed articles published between 2023 and 2025. The literature search was structured to cover both GenAI and management fields. Accordingly, the following subject area (TS) query was used in the WoS Advanced Search interface:

TS = ((“generative artificial intelligence” OR “generative AI” OR “ChatGPT”) AND (management OR “management research” OR “management studies” OR “strategic management” OR organization*)).

This search strategy has been expanded to encompass different uses of the concept of GenAI; relevant key concepts have been used together to ensure that studies involving management, strategic management, and organizational contexts are included in the analysis.

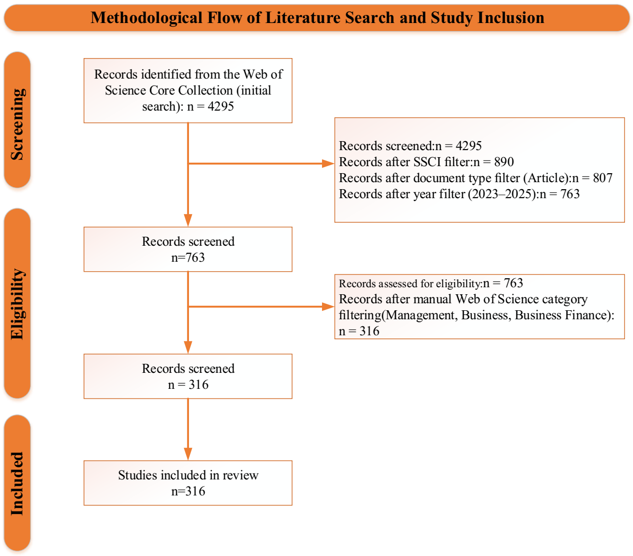
The initial dataset obtained as a result of the determined search strategy consisted of a total of 4295 records. In the first stage of the scanning process, the dataset was filtered within the scope of the SSCI, and as a result of this process, the number of records was determined to be 890. In the next stage, only publications of the article type were included in the analysis; as a result of this filtering, the number of records was reduced to 807.

In line with the study’s objective, the scope of analysis was limited to publications released between 2023 and 2025 to capture the literature explosion and most current academic output following the launch of ChatGPT in November 2022. 763 records meeting this criterion were evaluated. During the eligibility phase, a manual screening process was conducted based on WoS subject categories. In this process, studies not directly related to management, business, and business finance were excluded from the dataset. As a result of this screening process, 316 publications found to be consistent with the study’s aim and scope formed the final dataset.

The eligibility phase involved additional data filtering through WoS subject categories, which researchers used for manual screening of the dataset. Specifically, articles were retained if they were classified under at least one of the following WoS categories: Management, Business, Business Finance, or Economics. Articles classified only under Computer Science, Engineering, Education, or Medicine were excluded because they did not directly address management-related research questions. However, studies on management education were retained when they were also indexed under Management, Business, Business Finance, or Economics and directly addressed the managerial or organizational implications of GenAI. The evaluation process resulted in a dataset that contained research about GenAI from managerial, organizational, and strategic perspectives instead of technical or domain-specific studies. Two researchers performed manual screening independently while they discussed any discrepancies to achieve agreement about which studies to include.

The scanning, filtering, and selection process in question is structured in accordance with the PRISMA 2020 reporting principles and aims to ensure the transparency and reproducibility of the analysis process. As a result, the total number of publications evaluated in the bibliometric analyses conducted within the scope of this study and included in the quantitative synthesis was determined to be 316. From this dataset, the distribution of publications by year, citation structures at the author and country levels, co-authorship networks, and keyword co-occurrence patterns were analyzed. VOSviewer (version 1.6.20) software was used for data visualization and mapping. The study selection and filtering process is summarized in Figure 1 [23].

Figure 1
PRISMA 2020 flowchart



The temporal distribution of the 316 publications indicates a substantial increase in research output on GenAI and management. The number of publications rose from 18 in 2023 to 116 in 2024 and reached 182 in 2025, suggesting a rapidly expanding scholarly interest in this field. The analytical framework rests on three assumptions. First, bibliometric links are treated as observable proxies of intellectual structure, consistent with science-mapping theory. Second, the 2023–2025 SSCI corpus is used as the empirical boundary because ChatGPT created a distinct post-2022 research trajectory in management. Third, keywords are interpreted as thematic signals rather than direct measures of organizational effects; therefore, all main claims are supported with statistical validation, robustness checks, and null-model comparisons.

2.1. Computational pipeline and statistical analysis

To complement the descriptive bibliometric mapping, we implemented a Python-based computational pipeline for network analysis, topic modeling, inferential testing, community detection, and distributional analysis using the powerlaw library [24];

detailed package versions and configuration settings were retained for reproducibility in the supplementary methodological notes.

The screening of candidate records was performed independently by two reviewers, with disagreements resolved by discussion; inter-rater agreement was high (Cohen's $\kappa = 0.87$). The co-authorship graph was constructed at the article level, with authors as nodes and co-authorship instances as edges weighted by frequency. We characterized the network through global descriptors (density, number of components, average clustering coefficient, diameter, and average shortest-path length of the largest component) and through node-level centralities (degree, betweenness, and eigenvector). Community structure was identified using the Louvain algorithm, with modularity Q reported as a measure of partition quality. Author productivity was tested against Lotka's inverse-power law by a Kolmogorov–Smirnov (KS) goodness-of-fit test, and the citation distribution was compared between power law and lognormal forms via likelihood ratio testing following the procedure of Clauset et al.'s work [25].

The longitudinal growth of the literature was assessed on monthly publication counts using the Mann–Kendall trend test, with Sen's slope reported as effect size. To complement the VOSviewer-based keyword co-occurrence mapping, we fitted latent Dirichlet allocation (LDA) models on the abstracts in the corpus for $k \in \{4, \ldots, 10\}$ and selected $k = 7$ as the coherence-optimal topic solution by maximizing the Cv topic coherence score [26]. The LDA coherence value should be interpreted cautiously. Although the seven-topic solution yielded the highest Cv score among the tested models, the absolute value ($Cv = 0.304$) indicates moderate semantic separation. This is consistent with the early and rapidly evolving nature of GenAI management research, where themes such as adoption, governance, ethics, productivity, innovation, and human–AI collaboration often intersect within the same studies. Therefore, LDA was used not as stand-alone evidence of thematic structure but as a complementary semantic validation of the keyword co-occurrence and overlay analyses. Representative terms for each topic are provided in the supplementary appendix to support transparency and interpretability. Temporal shifts in keyword prevalence between the early sub-period (2023–2024) and the late sub-period (2025) were tested with two-sided Fisher's exact tests on 2×2 tables of keyword presence by period. Statistical significance was set at $\alpha = 0.05$.

To confirm these inferences statistically, four additional procedures were applied. Category-level thematic change was tested with two-sided Fisher's exact tests on predefined normative/-governance and performance/productivity keyword dictionaries across the early (2023–2024) and late (2025) sub-period, complemented by a McNemar test of overall category prevalence. Thematic diversification was assessed with a permutation-based rarefaction of unique author keywords (1000 draws), controlling for the larger number of articles in the late period. Author- and country-level associations among productivity, citation impact, and network connectivity were quantified with Spearman rank correlations. Finally, the observed Louvain modularity of the co-authorship network was benchmarked against a null distribution of 500 degree-preserving configuration-model graphs, yielding an empirical p-value and z-score. All tests used a 0.05 significance threshold and a fixed random seed (42) for reproducibility. To ensure reproducibility, the full workflow was implemented as a predefined sequence: WoS record export, SSCI/article/year filtering, manual eligibility screening, VOSviewer mapping, and Python-based statistical validation. The VOSviewer analyses used fractional counting, a minimum keyword occurrence threshold of two, and co-authorship inclusion based on at least one publication and one citation. Network maps were interpreted using total link strength, cluster membership, and largest-component structure. All Python scripts used fixed package versions and seed value 42; the same cleaned dataset was used across trend testing, Lotka-law assessment, citation-distribution fitting, LDA topic modeling, Fisher exact tests, Spearman correlations, rarefaction, and configuration-model robustness checks. The technical contributions of the proposed analytical framework are summarized in Table 1.

The statistical validation strategy and corresponding findings are presented in Table 2. These procedures were used to test the stability of the bibliometric inferences rather than to introduce a separate empirical experiment; therefore, the statistical strategy and its results are reported once in Tables 1 and 2 to avoid analytical redundancy.

The monthly publication count exhibited a strong and statistically significant upward trend across the observation window (Mann–Kendall $\tau = 0.570$, $Z = 4.42$, $p < 0.001$), with a Sen's slope of 0.50 articles per month, indicating a sustained rather than transient expansion of the literature following the public release of ChatGPT. Authorial productivity ($n = 1003$ unique authors generating 1097 authorship instances across 316 articles) was concentrated at the low-output end of the distribution: 926 authors (92.3%) contributed a single article, and only 14 authors contributed three or more. The observed productivity distribution was consistent with Lotka's inverse-power law ($\hat{\alpha} = 4.12$, KS $D = 0.010$, critical value at $\alpha = 0.05 = 0.043$), supporting the characterization of the field as one in early formation, where a small core of recurrent contributors has yet to consolidate.

The citation distribution was right-skewed and heavy-tailed (mean = 21.7, median = 8.0, maximum = 749 citations) but was significantly better explained by a lognormal than a power-law functional form (log-likelihood ratio $R = -7.31$, $p = 0.030$), suggesting that impact accrual in this corpus follows a multiplicative-growth rather than a strictly scale-free dynamic. The co-authorship network (979 nodes, 2077 edges, density = 0.0043) fragmented into 226 connected components, with the largest component containing only 72 authors (7.35% of the total). The average clustering coefficient was high (0.84), the diameter of the largest component was 8, and its average shortest-path length was 3.26, while Louvain community detection yielded a modularity of $Q = 0.951$ across 229 communities. Together, these descriptors indicate a topology of tight, locally cohesive teams that have not yet merged into an integrated collaboration core, a structural signature consistent with the emerging status of the field.

LDA on the abstracts identified $k = 7$ as the coherence-optimal topic solution ($Cv = 0.304$), with semantically interpretable themes corresponding to innovation and knowledge work, human–AI work design, service and supply-chain applications, strategy and automation, organizational adoption and performance, and management education. To improve the interpretability of the LDA output, the seven-topic solution was further examined through its most representative terms. As shown in Table 3, the topics indicate partially overlapping but analytically distinguishable domains, reflecting the multidimensional development of GenAI research in management.

Period-comparison analysis revealed a measurable thematic rotation between the early and late sub-periods: the term *ChatGPT* declined as a stand-alone keyword from 34.4% (33 of 96 articles) to 16.8% (37 of 220 articles, Fisher's exact $p = 0.001$), while the broader label *generative AI* expanded its share from 31.3% to 38.2% ($p = 0.25$). Author-keyword diversity increased

Table 1
Technical contributions of the proposed analytical framework

Analytical component	Method/algorithm	Output/measure	Function in the study
Bibliometric mapping	VOSviewer co-occurrence and co-authorship algorithms	Cluster maps and total link strengths	Identifies thematic and collaborative structures
Network analytics	Graph descriptors and node-level centralities (NetworkX 3.3)	Density, clustering, betweenness and eigenvector centrality	Quantifies the topology of the collaboration network
Community detection	Louvain modularity optimization	Modularity Q, community count	Identifies cohesive author groupings
Productivity distribution	Lotka's inverse-power law with Kolmogorov–Smirnov goodness-of-fit test	$\hat{\alpha}$, KS D	Tests the maturity of the authorial base
Citation distribution	Power law vs lognormal likelihood ratio testing [25]	R, p	Characterizes the form of impact accrual
Longitudinal trend	Mann–Kendall test with Sen's slope	τ, Z, slope	Tests monotonic growth of the literature
Topic modeling	Latent Dirichlet allocation with Cv coherence [26]	Optimal k, Cv score	Provides a semantic complement to keyword co-occurrence
Thematic shift testing	Two-sided Fisher's exact test on 2×2 period tables	Odds ratio, p	Validates the temporal rotation of themes
Null-model validation	Configuration-model randomization of the co-authorship graph (500 graphs)	Modularity z-score, empirical p	Tests whether the fragmented topology exceeds chance
Cross-level correlation	Spearman rank correlation at author and country levels	ρ, p	Tests whether productivity, impact and connectivity coincide
Diversity rarefaction	Permutation-based rarefied keyword richness (1000 draws)	Richness, permutation p	Tests thematic broadening net of sample-size growth

Table 2
Statistical validation strategy and findings

Research claim	Statistical procedure	Variable(s)	Result	Interpretation
Normative themes did not statistically displace performance themes at the keyword-category level	Two-sided Fisher's exact test on category presence × period	Normative vs performance keyword categories (early vs late)	Normative 20.9% → 18.4%, $p = 0.65$; performance 12.7% → 17.0%, $p = 0.42$	No significant category-level rotation; theme balance is stable while specific new constructs emerge
Thematic vocabulary expanded beyond what sample growth alone would produce	Rarefied keyword richness (permutation, 1000 draws)	Unique author keywords by period	Early 427; late 849; rarefied late ($n = 110$) $= 486 \pm 15$; $p = 0.001$	Genuine diversification rather than a sample-size artifact
The fragmented-yet-cohesive topology is structural, not a chance artifact	Observed Louvain modularity vs configuration-model null (500 graphs)	Modularity Q of the co-authorship network	$Q = 0.951$ vs null 0.483 ± 0.003; $z \approx 141$; empirical $p = 0.002$	High modularity exceeds random expectation; fragmentation is genuine
Publication productivity only partially coincides with collaboration centrality	Spearman rank correlation (author level)	Papers vs degree/eigenvector centrality ($n = 1030$)	ρ (degree) $= 0.31$; ρ (eigenvector) $= 0.27$ (both $p < 0.001$)	Weak-to-moderate link; high output does not guarantee a central network position

(*Continued*)

Table 2
(*Continued*)

Research claim	Statistical procedure	Variable(s)	Result	Interpretation
Publication count alone does not explain citation impact	Spearman rank correlation (author level)	Papers vs total citations (n = 1030)	$\rho = 0.20$ ($p < 0.001$)	Weak coupling; a few early conceptual works attract disproportionate citations
Country-level output, citation impact and connectivity are positively but imperfectly aligned	Spearman rank correlation among country indicators	Publications, citations, link strength (n = 166)	Pubs–links $\rho = 0.83$; pubs–citations $\rho = 0.62$; citations–links $\rho = 0.49$ (all $p < 0.001$)	Strong but imperfect coupling; rank reversals exist (high-impact, lower-output countries)

substantially across periods, from 427 to 849 unique author keywords, and constructs such as *human–AI collaboration, agentic AI*, and *employee engagement* appeared exclusively in the late sub-period, consistent with the diversification pattern observed in the overlay visualization.

To validate the inferences drawn from the descriptive maps, a set of confirmatory tests was conducted. First, the apparent thematic shift was tested directly at the level of predefined keyword categories. A normative/governance dictionary (ethics, governance, AI literacy, agentic AI, accountability, responsibility, trust, regulation, privacy, transparency, bias, fairness, safety, inclusion) and a performance/productivity dictionary (performance, productivity, efficiency, automation, competitiveness, cost, optimization, profit, growth) were applied to the author keywords, and category presence was compared across the early (2023–2024) and late (2025) sub-period using two-sided Fisher's exact tests. Neither category shifted significantly: normative themes moved from 20.9% to 18.4% ($p = 0.65$) and performance themes from 12.7% to 17.0% ($p = 0.42$), and the two categories did not differ significantly in overall prevalence across the corpus (McNemar $p = 0.24$). The thematic development therefore reflects a broadening of the field rather than a measurable displacement of performance themes by normative ones. This broadening interpretation is itself supported statistically: vocabulary richness expanded significantly even after rarefaction to equal article counts (early 427 vs rarefied late 486 $\pm$ 15 unique keywords; permutation $p = 0.001$), and specific normative constructs (agentic AI, AI literacy, human–AI collaboration, employee engagement) appeared exclusively in the late sub-period. Second, at the author level, publication productivity was only weakly to moderately associated with collaboration centrality (Spearman $\rho = 0.31$ for degree and $\rho = 0.27$ for eigenvector centrality, both $p < 0.001$) and weakly associated with total citations ($\rho = 0.20$, $p < 0.001$), confirming that output volume alone explains neither network position nor citation impact. Third, the fragmented-yet-cohesive structure of the collaboration network was confirmed against chance: the observed modularity ($Q = 0.951$) far exceeded that of 500 degree-preserving configuration-model graphs (null $Q = 0.483 \pm 0.003$; $z \approx 141$; empirical $p = 0.002$). Finally, at the country level, publication output, citation impact, and connectivity were positively but imperfectly correlated (publications–links $\rho = 0.83$;

Table 3
LDA topic solution and representative terms

Topic	Interpretive label	Representative terms
Topic 1	Innovation and knowledge work	Innovation, creativity, knowledge, capability, value creation, business model, organizational learning
Topic 2	Human–AI work design	Human–AI collaboration, employee, skills, trust, AI literacy, role design, decision-making
Topic 3	Service and customer applications	Service, customer experience, marketing, hospitality, tourism, engagement, chatbot
Topic 4	Supply-chain and operational coordination	Supply chain, risk, coordination, resilience, operations, automation, logistics
Topic 5	Strategy, governance, and accountability	Strategy, governance, ethics, accountability, responsibility, transparency, regulation
Topic 6	Organizational adoption and performance	Adoption, performance, productivity, efficiency, implementation, competitiveness, digital transformation
Topic 7	Management education and research practice	Education, learning, assessment, management education, academic practice, research methods, bibliometrics

publications–citations $\rho = 0.62$; citations–links $\rho = 0.49$; all $p < 0.001$), with notable rank reversals in which several lower-output countries attained disproportionate citation impact.

2.2. Computational robustness experiments

Because the study is bibliometric, the term "experimental analysis" is operationalized as a series of computational robustness experiments rather than human-subject experimentation. Four experiments were run. (i) Threshold-sensitivity experiment: keyword co-occurrence and co-authorship maps were regenerated at minimum-occurrence thresholds of 2, 3, and 5. (ii) Time-window experiment: all analyses were recomputed splitting the corpus into early (2023–2024) and late (2025) windows. (iii) Dictionary-robustness experiment: the normative and performance keyword categories were each tested with a narrow and a broad dictionary. (iv) Topic-model robustness experiment: LDA was fitted for $k = 4$ to 10 topics, and the coherence (Cv) score was used to select the model. Finally, a null-model experiment compared the observed network modularity against 500 degree-preserving random graphs. The outcomes of all experiments are reported in the Results (Table 4) and confirm that the principal findings are stable across analytical choices.

Taken together, these robustness checks indicate that the main bibliometric patterns are stable and that the observed growth, thematic diversification, and fragmented collaboration structure are unlikely to be artifacts of a single threshold, dictionary, time window, or topic-model specification.

2.3. Co-author analysis

The author collaboration network is weighted according to total connection strength (weights = total connection strength). In contrast, the VOSviewer co-authorship map was restricted to the largest visually interpretable connected component after applying publication and citation thresholds; therefore, the numerical values obtained from the full-network analysis and those represented in the visualized subnetwork are not directly identical. The difference between the number of unique authors and the number of co-authorship network nodes arises because author productivity was calculated on the full cleaned author set, whereas network statistics were calculated after applying graph-construction and visualization eligibility criteria. In co-authorship analysis, at least one publication and at least one citation criterion were used to reliably and meaningfully reveal the network structure. Accordingly, the fragment network maps are included in the largest connected member scope consisting of 48 authors. As a result of the analysis, the network structure was determined to be 5 clusters, 28 connections, and a total connection strength of 311.

Considering the number of publications, the most prolific authors in the literature include Corvello, V. (4), Kunz, W. H. (3), Wirtz, J. (3), Aw, E. C. (3), Cham, T. (3), Ooi, K. (3), Dwivedi, Y. K. (3), Chen, L. (3), and Kraus, S. (3). However, it is observed that some authors with high publication volume are not part of a co-authorship network. This indicates that these authors contribute to the literature mostly independently or through limited collaborative structures. Therefore, it can be said that productivity based on publication volume does not always coincide with

Table 4
Statistical and experimental validation of key bibliometric inferences

Research claim	Validation method	Variables	Statistical result	Manuscript implication
The literature is expanding rapidly	Mann–Kendall + Sen's slope	Monthly publication counts	$\tau = 0.570$, $Z = 4.42$, $p < 0.001$; slope = 0.50/mo	Growth is statistically significant, not visual artifact
The field is at an early, broadly distributed stage	Lotka's law + KS test	Author productivity distribution	$\hat{\alpha} = 4.12$; KS D = 0.010 < 0.043; 92.3% single-article authors	Productivity follows Lotka; dispersed authorship
Citation impact is concentrated in a few works	Power law vs lognormal Likelihood Ratio (LR) test	Citation distribution (n = 234)	$R = -7.31$, $p = 0.030$; lognormal better than power law	Impact concentrated in early conceptual studies
Normative themes emerged alongside performance themes	Fisher's exact + McNemar	Keyword categories × period	Normative $p = 0.65$; performance $p = 0.42$; McNemar $p = 0.24$	No statistical displacement; thematic broadening rather than replacement
Thematic vocabulary diversified	Rarefaction/permutation	Unique keywords by period	Early = 427; late = 849; rarefied late = 486 ± 15; $p = 0.001$	Diversity not merely a sample-size effect
Output does not equal influence/centrality	Spearman + bootstrap CI	Papers vs citations/centrality	$\rho = 0.20$ [0.14,0.25]; $\rho = 0.31$ [0.26,0.36]	Productivity, impact, connectivity interpreted separately
Collaboration network is fragmented	Louvain Q vs configuration-model null	Modularity Q of the co-authorship network	$Q = 0.951$ vs 0.483; $p = 0.002$	Fragmentation exceeds chance
Thematic structure is robust to thresholds	Threshold-sensitivity experiment	Top-30 keywords at occ. 2/3/5	Jaccard = 1.00	Clusters stable across thresholds
Keyword clusters are semantically coherent	LDA k-sweep ($k = 4$–10)	Topic coherence Cv	$k = 7$, $Cv = 0.304$ (reported)	Semantic validation of keyword clusters

central position within a collaborative network. The most cited authors in the final dataset are presented in Table 5.

Table 5
Top 10 Most cited authors

Author	Citation
Gunasekara, A.	640
Lim, W. M.	640
Pallant, J. I.	640
Pallant, J. L.	640
Pechenkina, E.	640
Kraus, S.	242
Mariani, M.	215
Dwivedi, Y. K.	213
Kunz, W. H.	184
Wirtz, J.	184

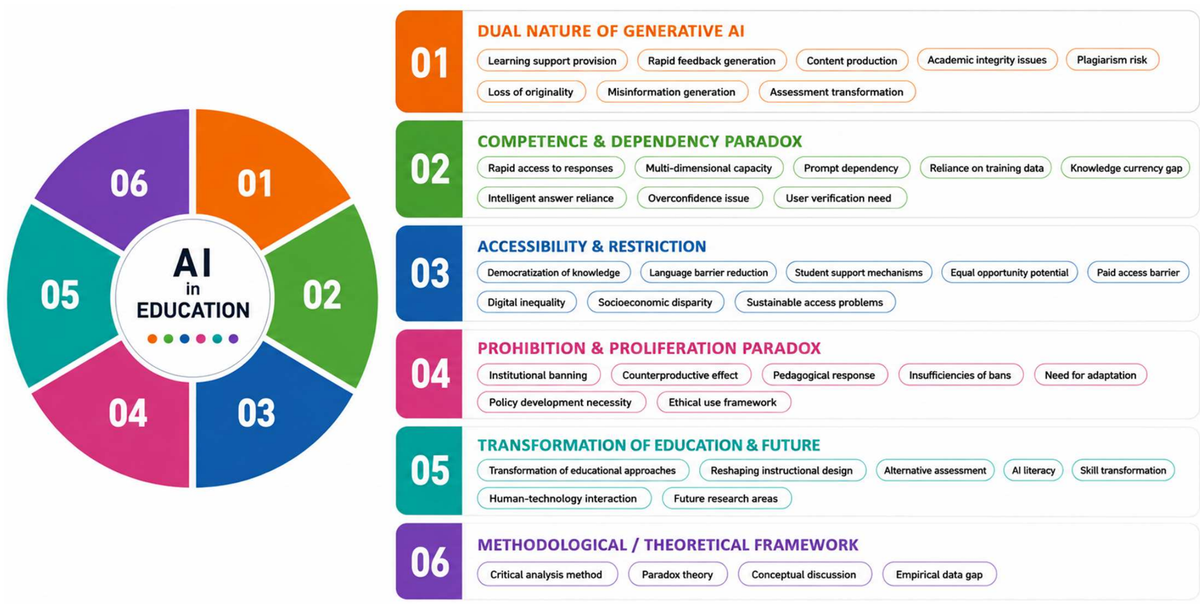
Table 5 shows the top 10 most cited authors in the management literature on GenAI. Gunasekara, A., Lim, W. M., Pallant, J. I., Pallant, J. L., and Pechenkina, E. rank first with 640 citations, followed by Kraus, S. (242), Mariani, M. (215), Dwivedi, Y. K. (213), Kunz, W. H. (184), and Wirtz, J. (184). These findings indicate that citation impact in this emerging field is not determined by publication volume alone.

Notably, the most cited contribution is the collaborative study by Gunasekara et al. [27], titled "Generative AI and the Future of Education: Ragnarök or Reformation? A Paradoxical Perspective from Management Educators." Its prominence suggests that early influence in this literature was shaped by studies that framed GenAI not only as an educational technology but also as a managerial and governance challenge involving decision-making, organizational adaptation, assessment integrity, and institutional policy. More broadly, the citation profile indicates that early influence is concentrated in a small group of highly cited, agenda-setting studies positioned at the intersection of management education, innovation management, service systems, digital transformation, and responsible governance. Citation impact therefore appears to reflect not only publication volume but also the capacity of selected studies to define the conceptual and managerial terms of the emerging debate. The deductive theme and code structure underlying these highly cited publication is presented in Figure 2.

3. Cited Publication

Citations from a variety of sources, including those found in GenAI-related academic research, as well as more general management-related sources, were reviewed in order to get a sense of the contribution to the field of management that Steve Kraus has made through his research. The number of citations to his work totaled 242, with the majority of those (63%) being related to just two of Kraus's papers. The paper that has received the most citations is his latest work, "On the use of AI-based tools like ChatGPT to support management research," including potential ways ChatGPT and similar LLMs could be used to support and potentially enhance management research including using them as a potential starting point for a systematic literature review. While the advantages of AI—such as speed, reproducibility, and objectivity—are emphasized, the necessity of human oversight is also specifically highlighted due to issues like lack of contextual sensitivity, generation of erroneous content, and ethical risks [28]. Additionally, 86 citations come from the study titled "The burgeoning role of literature review articles in management research: an introduction and outlook," which addresses the role of literature reviews in management research. While this study does not directly center on AI, it points to the integration of such technologies into systematic review processes and argues that literature reviews must evolve toward a more analytical and thematic structure [29]. In this regard, it can be said that Kraus's academic influence has been shaped through both AI applications and discussions on methodological transformation.

Figure 2
The deductive theme and code structure of a highly



Similarly, it is evident that Mariani has established a significant sphere of influence in the literature on GenAI and management. The author's citations are concentrated around two key studies. Of these, 149 are attributed to the study titled "Generative artificial intelligence in innovation management: A preview of future research developments." This study aims to establish a future research agenda at the intersection of GenAI and innovation management by combining a literature review with the Delphi method. The findings indicate that GenAI has the potential to create transformative effects in areas such as new product development, creativity, intellectual property, and organizational design [30]. On the other hand, 66 citations come from the study titled "Applications of generative AI and future organizational performance." This research analyzes the impact of GenAI usage on organizational performance and reveals that exploratory and exploitative innovation play a mediating role in this relationship. It is also noted that ethical dilemmas and environmental dynamics influence the process as moderating variables [31]. In this context, Mariani's work guides the literature through both theoretical and empirical contributions.

On the other hand, Dwivedi's academic influence is also evident in the GenAI literature. It is noteworthy that the total of 213 citations is concentrated around three key studies. A significant portion of these stems from the innovation management study conducted jointly with Mariani [30]. Furthermore, the study titled "Understanding the impact of ChatGPT on tourism and hospitality: Trends, prospects and research agenda" stands out with 47 citations. This study discusses how ChatGPT is transforming business processes in the tourism and hospitality sectors and presents a comprehensive research agenda [32]. Furthermore, the study titled "ChatGPT and service: Opportunities, challenges, and research directions" is also cited 17 times in the literature. This research addresses both the opportunities offered by GenAI in the context of service management and the associated ethical and operational risks [33]. In this context, Dwivedi's contributions demonstrate a multidimensional structure in terms of both sectoral applications and theoretical framework development.

Finally, the work of Kunz and Wirtz also holds a significant place in the literature on GenAI and management. The authors' combined 184 citations are concentrated around three key studies. Of these citations, 93 come from the study titled "Digital transformation: A multidisciplinary perspective and future research agenda" [34], which addresses digital transformation from a multidisciplinary perspective. This study comprehensively evaluates the impacts of AI and digital technologies on businesses and organizations. Additionally, the study titled "Corporate digital responsibility (CDR) in the age of AI: Implications for interactive marketing" stands out with 74 citations. This study addresses corporate responsibility within the framework of principles such as ethics, data security, and transparency in the use of AI and offers significant implications for management processes [35]. Furthermore, a study examining ChatGPT's role in the context of service management also contributes to the literature [33]. In this context, Kunz and Wirtz's studies offer a comprehensive contribution centered on digital transformation, ethical responsibility, and service management. The authors with the highest co-authorship connection strength are listed in Table 6.

According to the total linkage strength analysis, the top 10 most connected authors in the co-authorship network were determined as Greenwood, M. (29), Islam, G. (29), Kunz, W. H. (27), Wirtz, J. (27), Aw, E. C. (27), Cham, T. (27), Ooi, K. (27), Dwivedi, Y. K. (23), Sigala, M. (22) and Tan, G. W. (22).

Table 6
Top 10 authors with the highest co-authorship connection based on total connection strength

Author	Total connection power
Greenwood, M.	29
Islam, G.	29
Kunz, W. H.	27
Wirtz, J.	27
Aw, E. C.	27
Cham, T.	27
Ooi, K.	27
Dwivedi, Y. K.	23
Sigala, M.	22
Tan, G. W.	22

The co-authorship network visualization required VOSviewer to function with a minimum cluster size of 3 authors and a resolution parameter of 1.0 to achieve proper cluster detection and generate clean results. The LinLog/modularity layout algorithm optimized node positions through its algorithm, which placed authors with strong collaborative ties in proximity to each other in the visualization. The network underwent refinement through the selection of its biggest connected component, which aimed to study the most unified collaboration patterns in the field. The co-authorship network structure among the authors is visualized in Figure 3.

Figure 3 presents the co-authorship network, illustrating the collaboration structure among authors. Kunz, W. H., Dwivedi, Y. K., and Wirtz, J., who are among the top 10 authors in terms of publication volume and citation impact, also occupy a central position within the co-authorship network with high overall linkage strength values. Conversely, the absence of some authors with high overall linkage strength from the co-authorship map stems from the concentration of their collaborative relationships within subsets independent of the main collaboration network. Because only the most strongly connected component was considered in the mapping, authors who did not have a direct connection to the main collaboration network but had strong relationships within themselves were excluded from the visualization.

4. Country Citation Analysis

An international citation network was created to examine the citations received by publications according to their country of origin. The research analyzed nations that published at least one paper and received at least one citation reference; the study evaluated 52 countries based on these specific requirements. The network structure contains 11 clusters that connect through 393 links that produce a total link strength of 1077. The United States leads all countries with 1305 citations, followed by Australia with 1114 citations and the United Kingdom with 884 citations. The United States maintains the highest link strength total with 233, followed by the United Kingdom at 207 and the People's Republic of China at 194. Given that Figure 4 is intended to preserve the complete country-level citation topology, the network was not simplified by removing lower-link-strength countries. Instead, the interpretation focuses on the most central countries by total link strength, namely, the United States, the United Kingdom,

Figure 3
Co-authorship network showing collaboration among authors weighted by total connection strength

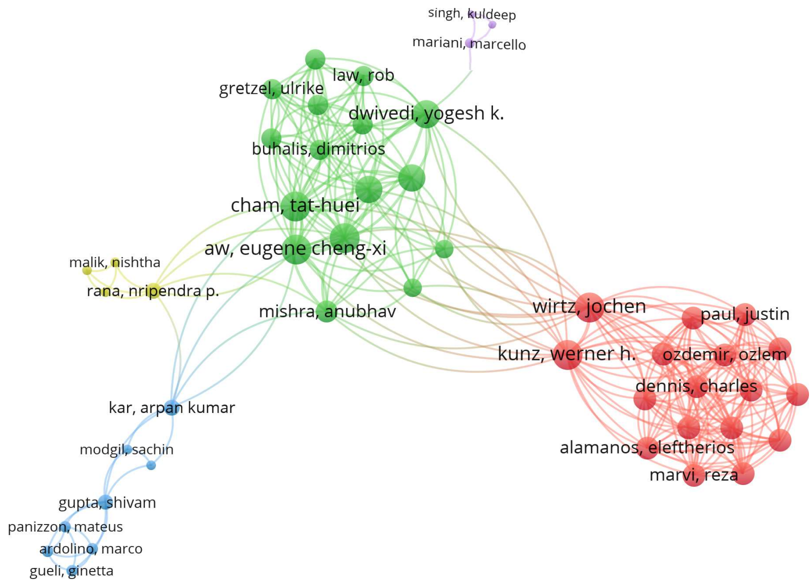
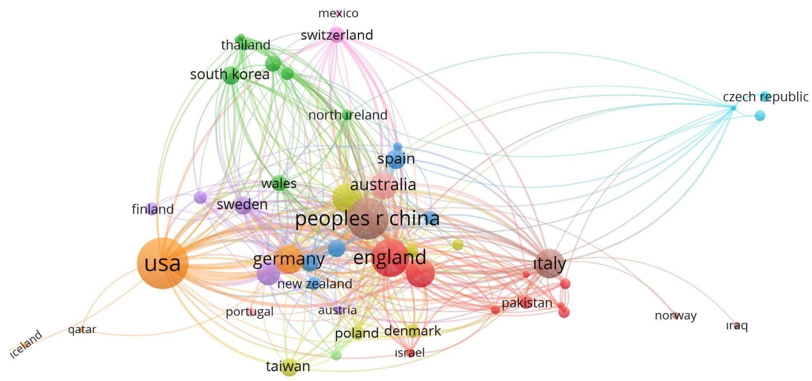


Figure 4
Citation network showing cross-country citation ties in generative artificial intelligence literature



and the People's Republic of China, which function as the principal citation-network hubs within the analyzed SSCI corpus. The research shows that these nations achieve both high citation rates and function as key nodes that link different parts of the worldwide citation network. The United States, together with the People's Republic of China and the United Kingdom, leads scientific research publication globally because they publish the highest number of scientific papers and articles. However, these country-level findings should be interpreted with caution because the SSCI-indexed WoS corpus may privilege English-language journals and internationally visible publication outlets. As a result, the observed dominance of particular countries reflects their position within the analyzed SSCI corpus rather than definitive evidence of global intellectual leadership across all languages, regions, and indexing systems. Generally speaking, indicators such as citation impact, total connectivity strength, and number of publications do not always perfectly correlate at the country level; some countries have high publication volumes, while others, despite having a more limited number of publications, exhibit stronger citation impact and network centrality. As shown in Figure 4, node sizes represent publication volume, while line thicknesses represent the intensity of citation interaction between countries.

These citation-based clusters should not be interpreted as direct evidence of thematic specialization. Rather, they indicate patterns of citation proximity, regional linkage, and intellectual connectedness among countries.

The clusters in Figure 4 display different nation groups through color, which indicates their participation in thematic or regional collaboration networks based on their citation patterns. The network structure shows multiple independent citation communities that exist as separate entities.

The orange cluster, which contains the United States at its center, represents the biggest network that exercises the most influence, while the United States functions as the main connection point that links various international research pathways. The red cluster contains England, Italy, Brazil, Turkey, and several Middle Eastern nations, which include the United Arab Emirates, Cyprus, and Pakistan. This cluster may reflect citation proximity among countries with shared or overlapping management research interests, but its thematic interpretation should be treated cautiously. The green cluster consists mainly of Southeast Asian nations, which have Malaysia and Thailand as their essential research centers, together with Wales and Northern Ireland for studying technology adoption and digital transformation. The yellow cluster contains Australia, the People's Republic of China, and South Korea, which form a Pacific Rim research network that may be associated with innovation-oriented and strategic AI-related research streams. The purple cluster contains Germany, Sweden, New Zealand, and Austria, which indicates their interest in governance and sustainable AI solutions and human-oriented AI systems. The pink/light purple cluster contains Switzerland and several Nordic countries because researchers in this group may be linked to governance-oriented and ethical AI discussions, although this interpretation requires keyword-level confirmation.

The visualization displays publication output through node sizes, which reveal that the United States has the largest node, followed by England, Australia, and the People's Republic of China, because these countries lead the development of GenAI management research. The thickness of lines between countries shows the strength of their citation links, which the United States uses to connect all regional groups through its dense orange connections. The strong connections between England and other European nations, as well as between Australia and Asian countries, reveal the presence of secondary hubs that facilitate regional knowledge diffusion. The field achieves worldwide integration through bridge nodes, which include Germany, Australia, and New Zealand because these countries exist at cluster boundaries.

4.1. Keyword analysis

In the keyword co-occurrence analysis, keywords that appeared at least twice were included in the analysis. Accordingly, 148 of the total 1184 keywords in the dataset met the determined threshold. To make the network structure more meaningful and interpretable, only the largest connected component was considered in the analyses; thus, 147 of the 148 keywords were included in the final analysis. As a result of the analysis performed on the 147 keywords that appeared at least twice and had a co-occurrence relationship, 20 clusters, 595 connections, and a total connection strength of 919 were identified in the network structure.

When the most frequently used keywords in the analyzed articles are evaluated, the phrase "generative artificial intelligence" ranks first with 201 occurrences. This is followed by "artificial intelligence" with 100 occurrences and "digital transformation" with 13 occurrences. In terms of total connection strength, "generative artificial intelligence" (362), "artificial intelligence" (233), and "ethics" (45) stand out as the keywords with the strongest thematic connections within the network. The keyword co-occurrence structure of the GenAI management literature is shown in Figure 5.

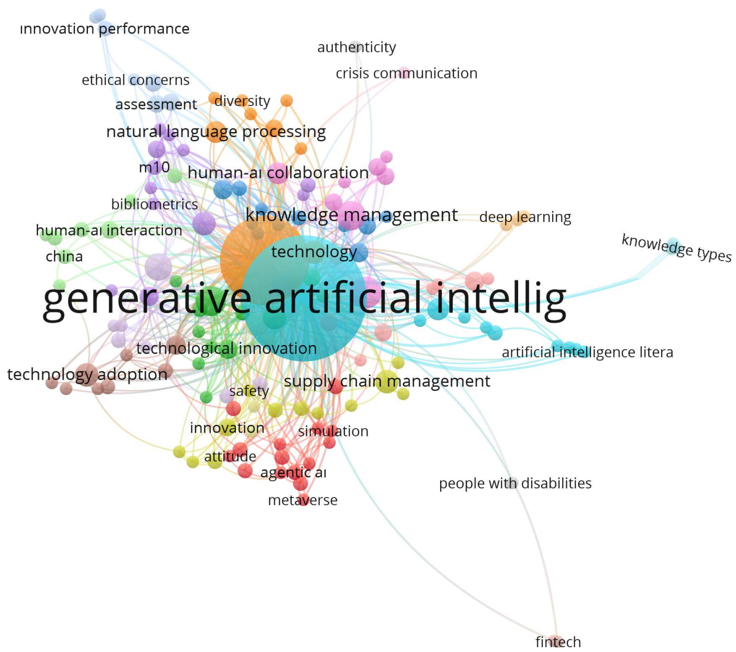
Overlay visualization analysis reveals the distribution of keywords according to their average publication year. Accordingly, keywords represented in yellow tones and corresponding to more recent publications are predominantly focused on the year 2025. In particular, concepts such as agentic AI, AI literacy, knowledge types, metaverse, and people with disabilities are among the most current and emerging themes in the literature.

4.2. Thematic cluster analysis

The keyword co-occurrence network shows 20 different thematic clusters, which demonstrate that GenAI research in management contains diverse and multiple intellectual approaches. Although the keyword co-occurrence algorithm produced 20 clusters, the 10 most substantively meaningful and interpretable clusters are discussed below.

The first cluster (Orange–Upper Center) consists of two elements that focus on technological basics and methodological approaches: The technical and methodological foundation of the field exists within this cluster, which includes chatbot, meta-analysis, NLP, and theory as its core elements. The field combines fundamental AI generation systems with conversational AI and NLP-based systems and analytical methods that

Figure 5
Keyword co-occurrence network



researchers use to study these systems. The cluster occupies a central location in the space that encircles AI because it operates as a fundamental research stream that supports all other thematic areas.

Cluster 2 (Purple/Violet–Central Left): Sociotechnical Integration and Ethical Dimensions: The cluster contains research terms that include human–AI collaboration, AI ethics, bibliometrics, and methodological classification codes (e.g., m15, ml6, m10, o33, m1). The study investigates human–AI system interactions, while developers must follow ethical rules and scientists need to work together using various research approaches. The cluster contains systematic reviews and mapping studies, which study the field's structural evolution by using bibliometric and methodological indicators and historical analysis.

Cluster 3 (Blue/Cyan–Right Side): Knowledge Management and Organizational Capabilities: This cluster contains terms that include knowledge management, knowledge sharing, absorptive capacity, business value, AI literacy, and innovation management. The research examines business AI implementation for knowledge management system improvement and flexible organizational framework development, which produces strategic business value. The research stream about organizational learning and capability development functions independently as a separate field that studies these topics through methods that link to the core research methodology.

Cluster 4 (Pink/Magenta–Center Right): Family Businesses and Contextual Applications: The research contains a particular smaller section that studies family businesses through anthropomorphism as its main focus. The research stream in this cluster investigates how family-owned enterprises adopt, understand, and implement GenAI technology within their specific organizational structures. The term appears in research to indicate scientists are creating advanced studies that analyze particular circumstances.

Cluster 5 (Red–Lower Center): Strategic Management and Emerging Technologies: The cluster contains six essential terms, which are agentic AI, metaverse, risk management, technology management, risk analysis, and dynamic capabilities. The research examines the most complex strategic management topics through its investigation of self-governing AI technology and digital space development and organizational abilities required to manage unpredictable technological changes and market shifts. The last period brought high-level concept introductions that prove this cluster operates as an active research field that shows strong expansion.

Cluster 6 (Yellow/Lime Green–Lower Center Left): Innovation and Technical Infrastructure: The cluster unites two main terms, which include LLMs and innovation, to connect technical architecture discussions with their resulting innovative outcomes. The research investigates how particular GenAI infrastructure systems that use LLM technology affect organizational innovation management.

Cluster 7 (Green–Far Left): Technology Adoption Frameworks: The cluster contains peripheral terms that include technology adoption, Technology, Organization, Environment (TOE) framework, China, OpenAI, companies, technological innovation, and affordances. The research applies an adoption-focused standard methodology that uses current models to analyze how organizations implement GenAI systems. Research now investigates how technology adoption patterns differ between countries and cultural environments through the implementation of geographic indicators.

Cluster 8 (Brown/Beige–Far Left): Adoption Barriers and Implementation Challenges: The research in this cluster studies the operational challenges of GenAI deployment through its limited number of terms, which concentrate on system capabilities and deployment settings.

Cluster 9 (Light Blue–Upper Left): Educational Applications: The research cluster about GenAI in education exists at the edge because it investigates AI implementation in academic environments through assessment and higher education as its fundamental concepts, which focus on teaching approaches and learning evaluation methods.

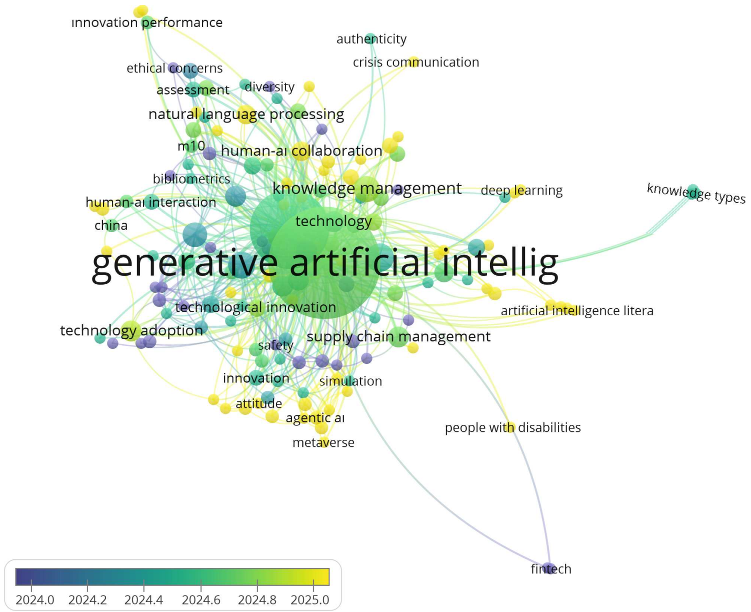
Cluster 10 (Gray/Cyan–Far Right and Bottom Right): Specialized and Emerging Applications: The research field includes three specific application domains that study isolated nodes involving people with disabilities, crisis communication, and supply chain management. The main discussion of the field does not fully incorporate these themes, yet they point to important research paths which scientists should investigate during future studies.

The research on GenAI in management shows diverse characteristics because it includes various theoretical frameworks, research methods, and practical uses. The core concept of AI connects various research domains because it exists at their common center. The central clusters show strong connection patterns because their technological–methodological (orange), sociotechnical–ethical (purple), and knowledge management (blue) clusters maintain their own distinct intellectual direction. The field expands into particular domains and actual environments through its peripheral clusters.

The human–AI collaboration cluster shows robust connections with the knowledge management cluster and the strategic management cluster, which serve as theoretical connectors to merge different concepts from the literature. The distinct clusters that study education, family businesses, and disability support programs demonstrate that these particular fields maintain their independence from the general academic dialog. The field shows its present organizational focus through its strategic examination of new concepts, which include agentic AI and the metaverse. The temporal evolution of keywords according to their average publication year is displayed in Figure 6.

In the overlay visualization (Figure 6), the color gradient represents the average publication year of articles associated with each keyword, ranging from dark blue/purple (2024.0) through green to bright yellow (2025.0), as indicated by the color scale bar at the bottom right of Figure 6. The research themes in GenAI management literature have undergone development according to the time-based analysis. The early-stage foundational themes that controlled the field during its initial development appear in dark blue and purple tones, which match the 2024.0–2024.2 period. The core idea of "generative artificial intelligence" and its connected terms "technology adoption" and "bibliometrics" appear in these darker tones, which shows these essential concepts started first and continue to exist throughout the entire research duration. The green-toned keywords from 2024.4 to 2024.6 represent the main topics that became more prominent during the development of the field. The research field expanded into particular application areas, which included organizational operations and specific domains, during mid-2024 through terms that focused on "knowledge management," "natural language processing," "higher education," and "supply chain management." Keywords in bright yellow tones (corresponding to 2024.8–2025.0) reflect the most recent and emerging themes in the literature. The terms "agentic AI," "skills," "people with disabilities," "innovation performance," "authenticity," and "knowledge types" stand out in bright yellow because they signify the most advanced research

Figure 6
Overlay visualization showing the distribution of keywords according to average publication year



findings at present. The visualization shows "knowledge types" as its most prominent section, which appears in cyan-yellow colors and exists independently from the primary research cluster. The visualization shows this category as an emerging field that scholars have just started to discuss in academic literature. The sequence of events demonstrates an obvious developmental sequence, which shows how the field first created basic technological principles during 2024.0–2024.2 before it started to study particular organizational uses of GenAI in different functional areas during 2024.4–2024.6, and finally, it focused on managing GenAI through normative, capability-based, and inclusive approaches during 2024.8–2025.0. The spatial distribution of yellow-toned keywords around the periphery of the network, rather than at its center, further confirms that the newest themes are expanding the field's boundaries rather than merely reinforcing its established core.

The small size of this subset indicates that agentic AI should be interpreted as an early emerging signal rather than a mature or generalizable research stream.

To translate the thematic signal of the bibliometric analysis into a concrete managerial implication, we conducted a focused content review of the subset of articles that explicitly invoke the construct of agentic AI. Candidate articles were identified through keyword and abstract screening, yielding nine articles (eight published in 2025 and one in 2024). This shift marks a movement from advisory GenAI systems that summarize, recommend, or draft content toward agentic systems that may initiate, select, or execute managerial actions in domains such as recruitment, supply-chain coordination, finance, and customer engagement. The temporal concentration of this subset a single article in 2024 followed by an eightfold increase in 2025 marks agentic AI as one of the most recently crystallized themes in the corpus.

Across the nine articles, four position agentic AI within marketing and business-to-business settings, where autonomous systems are reconceptualized as relational actors requiring trust-calibrated interaction design and revised customer-engagement strategies. Two articles in human-resource management address the substitution and re-design of recruitment and selection practices when autonomous agents intervene in candidate evaluation. One operations-oriented study treats agentic AI as a structural component of large-scale supply-chain orchestration. A finance article examines institutional asset management under autonomous decision systems, and a commentary article in the management-theory literature addresses the conceptual implications of machine agency for organizational accountability.

Three managerial concerns recur across these articles: (i) delegation boundaries which decisions, and over which time horizons, may legitimately be ceded to an autonomous system without erosion of human accountability; (ii) monitoring infrastructure how organizations audit non-human action in the absence of direct behavioral observation; and (iii) role re-design how human workflows, performance metrics and reporting lines adapt when AI systems act rather than merely advise. The convergence of these concerns across functional domains suggests that governance frameworks developed for traditional, advisory AI may be insufficient for agentic deployments, in which the locus of decision-making, not merely its informational input, partly shifts to the system. Building on this convergence, we propose a three-tier control structure for organizations adopting agentic AI: (1) ex ante constraints, encoded as scope, authority, and policy boundaries at deployment; (2) in-flight monitoring, through structured decision-logging, anomaly thresholds, and routine escalation procedures; and (3) ex post review, with human-in-the-loop oversight triggered when outcomes exceed predefined risk envelopes. This framework operationalizes the normative themes of ethics, AI literacy, and accountability that the keyword and topic analyses identify as the emerging core of post-ChatGPT management scholarship.

Table 7 shows that the corpus is methodologically diverse. Survey/SEM studies (22.8%) and conceptual/review papers (22.5%) dominate, while experiment/quasi-experiment studies (10.4%) and LLM performance/prediction analyses (11.1%) indicate the presence of empirical and computational validation within the field.

Table 7
Methodological profile of the 316 analyzed studies

Study type	Count	Share (%)
Survey/SEM/Partial Least Squares Structural Equation Modeling (PLS-SEM)	72	22.8
Conceptual/review	71	22.5
Qualitative/case study	48	15.2
Other/unclassified	48	15.2
LLM performance/prediction	35	11.1
Experiment/quasi-experiment	33	10.4
Bibliometric/scientometric	6	1.9
Secondary data/archival	3	0.9

The findings suggest that post-ChatGPT management research is shifting from whether GenAI can improve organizational efficiency to how its use can remain accountable, auditable, and strategically governed. For practice, this implies that GenAI adoption should be accompanied by role-specific AI literacy, explicit delegation limits, traceable decision records, and escalation rules for high-risk AI-mediated outputs.

5. Future Research Directions

The research results from this study revealed multiple critical research directions that scientists need to investigate. Organizations need to create governance systems that will track AI autonomy because GenAI systems now function as semi-autonomous or autonomous decision-making agents. Research needs to study Managerial practices when dealing with human–machine decision authority restrictions and determine which organizational systems and coordination methods develop when AI performs essential mental tasks and how staff members respond to AI decision outputs.

Research about organizational dimensions of AI literacy, which serves as a vital factor for AI adoption success, remains scarce despite widespread recognition of its importance. Research in the future needs to determine which cognitive abilities, technical skills, and ethical knowledge managers require to develop AI literacy and how organizations should implement training programs for different employee demographics. Research needs to prove how AI literacy affects AI program success rates and how it enables organizations to link AI system deployment to their business results.

The existing research lacks sufficient development of ethical considerations, which form a vital yet underdeveloped field of study. The current discussion about ethical issues lacks sufficient methods that organizations need to implement for GenAI control. Research in the future needs to determine which governance systems and control mechanisms work best for AI responsibility, how organizations achieve balance between operational efficiency and ethical, legal, and social requirements, and how various organizational groups including executives, employees, customers, and regulators shape AI ethics policy development and implementation.

The existing research data consists mostly of cross-sectional and conceptual studies, which need longitudinal research approaches to analyze the time-based development of GenAI adoption. Research conducted over multiple years would reveal important information about how organizations develop their structures and cultures and work methods when GenAI systems start performing routine tasks. The research would show how long-term AI implementation impacts worker abilities and work content and professional development and workplace contentment. The research would determine if the expected advantages from AI implementation actually occur during the extended period of implementation. The research would identify which organizational elements, working with leadership approaches, produce successful project delivery.

Research should concentrate on studying how business operations change based on their industrial sector and surrounding conditions. The current body of research about GenAI focuses on general concepts but fails to recognize how different sectors, organizational sizes, and cultural settings affect its implementation. Research studies that compare different situations would demonstrate how healthcare, finance, and manufacturing sectors handle their ethical challenges, which produce unique results for their operational performance and strategic growth. The research needs to study how small businesses use AI systems differently than large corporations and how organizational GenAI management practices differ between cultures and under different regulatory systems.

6. Conclusion

This study aims to reveal the structural and thematic characteristics of the rapidly expanding GenAI literature in management, which emerged with the launch of ChatGPT, through bibliometric analysis. The analyses, conducted on 316 articles indexed in the SSCI database within the Web of Science Core Collection and published between 2023 and 2025, provide significant findings regarding the development dynamics of the literature in line with the research questions.

The findings show that GenAI research in management is expanding steadily rather than episodically. Monthly publication growth is significant, while the author base remains immature: 92.3% of authors contributed only one article, indicating a field still organized around a small emerging core. Thematic evidence further shows that ethics, governance, AI literacy, and agentic AI have not displaced productivity and performance concerns; instead, they have expanded the agenda alongside them. Collaboration patterns remain fragmented, with cohesive local teams yet to consolidate into a unified global network. At the country level, the United States, the United Kingdom, and the People's Republic of China occupy central positions in citation impact and knowledge connectivity.

The emergence of "agentic AI" in 2025 should be interpreted as an early theoretical signal rather than a settled paradigm shift. It indicates that management research is beginning to examine how autonomous GenAI systems may reshape accountability, oversight, and role design across organizational settings.

Implications for theory. The findings position GenAI as an emerging management research domain in which productivity-oriented debates now coexist with questions of agency, accountability, governance, and human–AI role design.

Implications for management practice. The proposed governance logic translates the bibliometric evidence into a practical managerial architecture by linking emerging theoretical concerns about agency, accountability, and autonomy with operational control mechanisms. Specifically, the three-tier structure of ex ante constraints, in-flight monitoring, and ex post review provides organizations with a scalable framework for governing GenAI systems that increasingly move from advisory support toward semi-autonomous or agentic decision-making roles.

Managers should treat GenAI adoption as an organizational governance challenge, requiring role-specific AI literacy, explicit delegation boundaries, auditable decision records, and escalation mechanisms for high-risk AI-supported decisions.

This study should be evaluated within the framework of some limitations. Taken together, the study provides a statistically validated map of an emerging field and offers a practical compass for managers and scholars. Its central implication is that the GenAI challenge in management is not only technological but also concerns governance capacity, ethical accountability, and the re-design of human roles around increasingly autonomous systems. First, the analyses are limited to SSCI-indexed articles in the Web of Science Core Collection. This single-source design strengthened metadata consistency, citation-link reliability, and reproducibility, but it may have excluded relevant studies indexed only in Scopus, Dimensions, Google Scholar, or regional databases. Therefore, the findings should be interpreted as a statistically validated map of the SSCI-indexed management literature rather than as a complete representation of all global scholarship on GenAI and management. Future studies should validate the present results using merged WoS–Scopus datasets and broader cross-database corpora. In addition, the SSCI-based design may underrepresent high-quality non-English or non-SSCI scholarship, including Chinese- and German-language studies; future research should therefore compare SSCI evidence with regional and multilingual databases to obtain a more globally balanced map. This may have led to the exclusion of studies in other major databases such as Scopus and earlier publications from the scope of the analysis. Second, the bibliometric analysis method reveals the structural and thematic characteristics of the literature through quantitative indicators; it does not offer a direct assessment of the content depth, theoretical frameworks, and methodological qualities of the studies. Furthermore, since keyword-based analyses depend on the terms used by the authors and the indexing processes in the database, it is possible that some conceptual overlaps or newly emerging themes may not be fully captured. Finally, the focus of the analyses on the post-ChatGPT period leads to an evaluation of the literature at an early stage of development; this results in limited observation of long-term trends and potential paradigm shifts.

Funding Support

This work was supported by the İİBF.25.02 project fund provided by the Scientific Research Projects Coordination Unit of Firat University.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Author Contribution Statement

Irem Tanyıldızı Baydili: Conceptualization, Methodology, Formal analysis, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Burak Tasci:** Software, Writing–original draft, Writing – review & editing, Visualization. **Sengul Dogan:** Validation, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Turker Tuncer:** Validation, Investigation, Writing – original draft, Writing – review & editing, Visualization.

References

[1] Boussioux, L., Lane, J. N., Zhang, M., Jacimovic, V., & Lakhani, K. R. (2024). The crowdless future? Generative AI and creative problem-solving. *Organization Science*, *35*(5), 1589–1607. https://doi.org/10.1287/orsc.2023.18430

[2] Bi, Q. (2023). Analysis of the application of generative AI in business management. *Advances in Economics and Management Research*, *6*(1), 36–41. https://doi.org/10.56028/aemr.6.1.36.2023

[3] Korzynski, P., Mazurek, G., Altmann, A., Ejdys, J., Kazlauskaite, R., Paliszkiewicz, J., . . . , & Ziemba, E. (2023). Generative artificial intelligence as a new context for management theories: Analysis of ChatGPT. *Central European Management Journal*, *31*(1), 3–13. https://doi.org/10.1108/CEMJ-02-2023-0091

[4] De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E., & Rizzo, C. (2023). ChatGPT and the rise of large language models: The new AI-driven infodemic threat in public health. *Frontiers in Public Health*, *11*, 1–8. https://doi.org/10.3389/fpubh.2023.1166120

[5] Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., . . . , & Wright, R. (2023). Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 1–63. https://doi.org/10.1016/j.ijinfomgt.2023.102642

[6] Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., . . . , & Fiedel, N. (2023). Palm: Scaling language modeling with pathways. *Journal of Machine Learning Research*, *24*(240), 1–113.

[7] Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., . . . , & Sutskever, I. (2021). Zero-shot text-to-image generation. In *International Conference on Machine Learning*, 8821–8831.

[8] Hartmann, J., Exner, Y., & Domdey, S. (2025). The power of generative marketing: Can generative AI create superhuman visual marketing content? *International Journal of Research in Marketing*, *42*(1), 13–31. https://doi.org/10.1016/j.ijresmar.2024.09.002

[9] Holmström, J., & Carroll, N. (2024). How organizations can innovate with generative AI. *Business Horizons*, *68*(5), 559–573. https://doi.org/10.1016/j.bushor.2024.02.010

[10] Alijoyo, F. A., Reddy, L. C. S., Selvi, V., Murugan, R., Kakad, S., & Balakumar, A. (2024). Transforming business with generative AI models applications trends. In *2024 International Conference on Artificial Intelligence and Quantum Computation-Based Sensor Application*, 1–5. https://doi.org/10.1109/ICAIQSA64000.2024.10882293

[11] Cui, Y. G., van Esch, P., & Phelan, S. (2024). How to build a competitive advantage for your brand using generative AI. *Business Horizons*, *67*(5), 583–594. https://doi.org/10.1016/j.bushor.2024.05.003
[12] Ratten, V., & Jones, P. (2023). Generative artificial intelligence (ChatGPT): Implications for management educators. *The International Journal of Management Education*, *21*(3), 100857. https://doi.org/10.1016/j.ijme.2023.100857
[13] Chowdhury, S., Budhwar, P., & Wood, G. (2024). Generative artificial intelligence in business: Towards a strategic human resource management framework. *British Journal of Management*, *35*(4), 1680–1691. https://doi.org/10.1111/1467-8551.12824
[14] Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, *66*, 111–126. https://doi.org/10.1007/s12599-023-00834-7
[15] Kane, T. B. (2025). Philosophical approaches to managing generative AI agents as artificial persons at work. *Journal of Business Analytics*, *8*(4), 250–266. https://doi.org/10.1080/2573234X.2025.2482652
[16] Sidra, S., & Mason, C. (2026). Generative AI in human-AI collaboration: Validation of the collaborative AI literacy and collaborative AI metacognition scales for effective use. *International Journal of Human–Computer Interaction*, *42*(7), 5084–5108. https://doi.org/10.1080/10447318.2025.2543997
[17] Zheng, X., Gildea, E., Chai, S., Zhang, T., & Wang, S. (2023). Data science in finance: Challenges and opportunities. *AI*, *5*(1), 55–71. https://doi.org/10.3390/ai5010004
[18] Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, *133*, 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070
[19] Klarin, A. (2024). How to conduct a bibliometric content analysis: Guidelines and contributions of content co-occurrence or co-word literature reviews. *International Journal of Consumer Studies*, *48*(2), 1–20. https://doi.org/10.1111/ijcs.13031
[20] Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods*, *18*(3), 429–472. https://doi.org/10.1177/1094428114562629
[21] Yeung, A. W. K. (2023). A revisit to the specification of sub-datasets and corresponding coverage timespans when using Web of Science Core Collection. *Heliyon*, *9*(11), e21527. https://doi.org/10.1016/j.heliyon.2023.e21527
[22] Lim, W. M., Kumar, S., & Donthu, N. (2024). How to combine and clean bibliometric data and use bibliometric tools synergistically: Guidelines using metaverse research. *Journal of Business Research*, *182*, 1–31. https://doi.org/10.1016/j.jbusres.2024.114760
[23] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *The BMJ*, *372*. https://doi.org/10.1136/bmj.n71
[24] Alstott, J., Bullmore, E., & Plenz, D. (2014). Powerlaw: A Python package for analysis of heavy-tailed distributions. *PloS One*, *9*(1), e85777. https://doi.org/10.1371/journal.pone.0095816
[25] Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*(4), 661–703. https://doi.org/10.1137/070710111
[26] Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, 399–408. https://doi.org/10.1145/2684822.2685324
[27] Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, *21*(2), 100790. https://doi.org/10.1016/j.ijme.2023.100790
[28] Burger, B., Kanbach, D. K., Kraus, S., Breier, M., & Corvello, V. (2023). On the use of AI-based tools like ChatGPT to support management research. *European Journal of Innovation Management*, *26*(7), 233–241. https://doi.org/10.1108/EJIM-02-2023-0156
[29] Kraus, S., Bouncken, R. B., & Yela Aránega, A. (2024). The burgeoning role of literature review articles in management research: An introduction and outlook. *Review of Managerial Science*, *18*(2), 299–314. https://doi.org/10.1007/s11846-024-00729-1
[30] Mariani, M., & Dwivedi, Y. K. (2024). Generative artificial intelligence in innovation management: A preview of future research developments. *Journal of Business Research*, *175*, 114542. https://doi.org/10.1016/j.jbusres.2024.114542
[31] Singh, K., Chatterjee, S., & Mariani, M. (2024). Applications of generative AI and future organizational performance: The mediating role of explorative and exploitative innovation and the moderating role of ethical dilemmas and environmental dynamism. *Technovation*, *133*, 103021. https://doi.org/10.1016/j.technovation.2024.103021
[32] Sigala, M., Ooi, K. B., Tan, G. W. H., Aw, E. C. X., Buhalis, D., Cham, T. H., . . . , & Ye, I. H. (2024). Understanding the impact of ChatGPT on tourism and hospitality: Trends, prospects and research agenda. *Journal of Hospitality and Tourism Management*, *60*, 384–390. https://doi.org/10.1016/j.jhtm.2024.08.004
[33] Sigala, M., Ooi, K. B., Tan, G. W. H., Aw, E. C. X., Cham, T. H., Dwivedi, Y. K., . . . , & Wirtz, J. (2024). ChatGPT and service: Opportunities, challenges, and research directions. *Journal of Service Theory and Practice*, *34*(5), 726–737. https://doi.org/10.1108/JSTP-11-2023-0292
[34] Paul, J., Ueno, A., Dennis, C., Alamanos, E., Curtis, L., Foroudi, P., . . . , & Wirtz, J. (2024). Digital transformation: A multidisciplinary perspective and future research agenda. *International Journal of Consumer Studies*, *48*(2), e13015. https://doi.org/10.1111/ijcs.13015
[35] Kunz, W. H., & Wirtz, J. (2024). Corporate digital responsibility (CDR) in the age of AI: Implications for interactive marketing. *Journal of Research in Interactive Marketing*, *18*(1), 31–37. https://doi.org/10.1108/JRIM-06-2023-0176

How to Cite: Baydili, I. T., Tasci, B., Dogan, S., & Tuncer, T. (2026). Generative AI in Management Research After ChatGPT: A Bibliometric Map of Themes, Influence, and Governance. *Artificial Intelligence and Applications*. https://doi.org/10.47852/bonviewAIA620210405