

Mapping the Intellectual Structure of Generative AI Research in Retail and E-commerce: A Bibliometric Review

Huynh Minh Khoi¹, Tran Thien Phuc²

¹SNB Retail And Services Joint Stock Company

²Ho Chi Minh City University of Foreign Languages – Information Technology

Received: 15.02.2026 | Accepted: 04.05.2026 | Published: 07.05.2026

*Corresponding author: Trần Thiên Phúc

DOI: [10.5281/zenodo.20072313](https://doi.org/10.5281/zenodo.20072313)

Abstract

Original Research Article

Generative artificial intelligence (GenAI) has rapidly transitioned from a frontier laboratory technology into a mainstream commercial layer of retail and e-commerce, reshaping personalization, recommendation, content generation, and conversational commerce. Despite an explosion of empirical and conceptual work since late 2022, the intellectual architecture of this body of knowledge remains fragmented. This study aims to map the intellectual structure, thematic clusters, and emerging research frontiers of GenAI scholarship in retail and e-commerce, with particular attention to agentic AI and retrieval-augmented generation (RAG) as emerging concepts and to country-level impact differentials. A PRISMA-guided bibliometric review was conducted on 445 Scopus-indexed articles published between 2023 and April 2026 (5,460 cumulative citations, 75 contributing countries). Seven complementary VOSviewer analyses were performed: keyword co-occurrence with curated thesaurus harmonization (26 mappings), co-citation of sources, co-citation of first authors, bibliographic coupling of documents, bibliographic coupling of journals, co-authorship of authors using Scopus identifiers, and co-authorship of countries. Full counting and association-strength normalization were applied throughout. The corpus exhibits exponential growth (12→90→218 articles in 2023–2025) and is structured around five thematic clusters: GenAI core and customer outcomes, AI chatbots and adoption, NLP/ML methods, e-commerce and customer experience, and large language models with recommender systems. Two parallel intellectual traditions anchor the field: a technical-AI tradition (Vaswani, Devlin, Brown, Radford, Touvron) and an adoption-behavioral tradition (Davis, Venkatesh, Ajzen, Hair, Dwivedi, Mogaji). Emerging concepts of agentic AI, retrieval-augmented generation, and conversational AI (average publication years 2025.3–2025.7) constitute the research frontier. Country impact is highly heterogeneous: China leads in volume (112 papers, 7.48 citations per paper) but the United Kingdom, India, the United Arab Emirates, Romania, Taiwan, and France produce small-but-impactful portfolios with citation per paper ratios ranging from 29 to 89. This is the first comprehensive bibliometric mapping of GenAI scholarship at the retail/e-commerce intersection, the first to document the dual technical-behavioral intellectual structure of the field, and the first to position agentic AI and RAG as emerging research frontiers using empirical co-occurrence overlay temporal evidence.

Keywords: Generative artificial intelligence, Large language models, Retail, E-commerce, VOSviewer, Agentic AI, Retrieval-augmented generation, Conversational commerce.

Copyright © 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

1. Introduction

The release of ChatGPT in November 2022 marked a discontinuity in the commercial adoption of artificial intelligence (AI), and retail and electronic commerce (e-commerce) have emerged as the most data-rich and consumer-facing testbeds for the new wave of generative AI (GenAI) capabilities (Hermann & Puntoni, 2024; Mogaji et al., 2024). Industry tracking services estimate that the global generative AI market grew from approximately USD 14.6 billion in 2022 to USD 45.6 billion in 2024 and is projected to exceed USD 207 billion by 2030, with retail, e-commerce, and digital marketing identified as the top three commercial verticals by deployment intensity (Cillo & Rubera, 2025; Kar et al., 2023). Within this commercial layer, large language models (LLMs), retrieval-augmented generation (RAG), and increasingly autonomous "agentic" architectures are reshaping core e-commerce primitives such as product search and discovery, conversational commerce, dynamic content generation, fake-review detection, recommender systems, and post-purchase service recovery (Roumeliotis et al., 2024; Wang et al., 2024; Zhao et al., 2024).

The corresponding research stream has expanded with comparable velocity. Our analysis of Scopus-indexed publications shows that articles addressing generative AI in retail or e-commerce contexts grew from 12 in 2023 to 90 in 2024 and 218 in 2025, with 125 additional articles published in the first quarter of 2026 alone — an effective doubling roughly every 12 months. Cumulatively, the corpus now consists of 445 peer-reviewed articles authored by researchers from 75 countries and has accumulated 5,460 citations. This explosive growth, combined with the intrinsically interdisciplinary character of GenAI research (spanning computer science, marketing, information systems, consumer psychology, and operations management), creates a pressing need for systematic intellectual mapping (Mariani et al., 2022; Vlačić et al., 2021).

Despite the proliferation, four interrelated gaps persist in the literature. First, prior bibliometric reviews of AI in marketing and commerce predate the generative-AI inflection or treat

GenAI as a marginal sub-topic embedded in broader AI surveys. Vlačić et al. (2021) reviewed AI in marketing through 2020, before the public emergence of ChatGPT. Mariani et al. (2022) covered AI in marketing, consumer research, and psychology with a similar pre-GenAI cut-off. Kumar et al. (2025) recently mapped AI chatbot research in marketing across 1996–2023 (242 articles) but explicitly bounded the inquiry to chatbot-as-conversational-agent and excluded most non-chatbot generative use cases such as creative content, product description generation, programmatic ad copy, RAG-based search, and agentic shopping assistants. Kar et al. (2023) provided an early exploratory review of generative AI in industrial applications without retail/e-commerce focus. No bibliometric review has yet specifically mapped the post-2022 generative-AI corpus at the retail/e-commerce intersection.

Second, the dual intellectual heritage of the field — in which technical advances from the deep-learning community (Vaswani et al., 2017; Devlin et al., 2019; Brown et al., 2020; Touvron et al., 2023) coexist with adoption-behavioral models from the consumer-research and information-systems traditions (Davis, 1989; Venkatesh et al., 2003; Ajzen, 1991; Dwivedi et al., 2023) — has not been quantitatively documented. Practitioners and reviewers routinely report a sense that "two communities" are working in parallel on overlapping problems, but evidence of this dual structure has not been provided through co-citation analysis on a representative corpus.

Third, the research frontier in retail/e-commerce GenAI is moving rapidly beyond first-generation chatbot and content-generation use cases toward more architecturally sophisticated paradigms: retrieval-augmented generation (RAG) for grounded, citation-aware product answers; agentic AI for autonomous shopping assistants and supply-chain co-pilots; multi-agent orchestration; and graph-based LLMs for knowledge-intensive retail reasoning (Jin et al., 2024). These emerging concepts are visible to practitioners but have not yet been systematically positioned within the bibliometric structure of the field.

Fourth, the geographic distribution of research output and impact in this nascent field is poorly understood. Whereas China has rapidly captured volume leadership (publishing 112 articles, or 25.2% of the corpus), preliminary indicators suggest that countries such as India, the United Kingdom, the United Arab Emirates, Taiwan, and France are producing portfolios with substantially higher per-article impact, raising important questions about the geography of intellectual leadership and collaboration patterns (Dwivedi et al., 2023; Kshetri et al., 2024).

Motivated by these gaps, the present study addresses six research questions:

RQ1. What is the publication trajectory and structural composition of the GenAI-in-retail/e-commerce research corpus from 2023 to early 2026?

RQ2. Which journals, authors, institutions, and countries are leading the conversation, and how does their productivity relate to scholarly impact?

RQ3. What is the intellectual structure of the field as revealed by co-citation of sources and first authors?

RQ4. How do bibliographic coupling and co-authorship reveal the emerging collaboration architecture of the field?

RQ5. What research gaps and high-leverage future directions can be identified for both academic and managerial audiences?

This study contributes to the literature in four ways. Theoretically, it provides the first comprehensive bibliometric mapping of GenAI in retail/e-commerce (2023–2026), documents the dual technical-behavioral intellectual structure of the field, positions agentic AI and RAG as empirically grounded research frontiers, and offers a typology of country impact patterns. Methodologically, it deploys seven complementary VOSviewer analyses with curated thesaurus harmonization (26 mappings) and overlay temporal visualization on a single, internally consistent corpus — a level of analytical depth rarely seen in marketing-bibliometric work. Practically, the study yields a journal portfolio map for researchers, a collaboration map for institutions, and a

capability map for retailers and platform operators. Policy-wise, it surfaces an emerging country-impact differential that is directly relevant to research-funding strategy in countries seeking to build generative-AI scholarly capacity.

2. Literature review

2.1 From predictive AI to generative AI in commerce

Artificial intelligence has been deployed in marketing and e-commerce for more than two decades, but the dominant paradigm until 2022 was *predictive* AI — supervised models for churn prediction, recommender ranking, demand forecasting, and click-through-rate optimization (Guha et al., 2021; Vlačić et al., 2021). The transformer architecture introduced by Vaswani et al. (2017) and progressively scaled into BERT (Devlin et al., 2019), GPT-2/3 (Radford et al., 2019; Brown et al., 2020), and the LLaMA family (Touvron et al., 2023) reframed the technological possibility frontier from prediction to *generation* — the synthesis of plausible text, images, code, and decisions conditioned on natural-language prompts. Hermann and Puntoni (2024) describe this transition as a movement "from predictive to generative AI" with profound implications for how firms design consumer-facing experiences. Cillo and Rubera (2025) provide a roadmap for innovation and marketing scholarship in the new paradigm, while Kar et al. (2023) catalog early industrial applications across both scientific and grey literature.

Generative AI use cases in retail and e-commerce now span a wide spectrum: programmatic generation of product titles, descriptions, and advertising copy (Brüns & Meißner, 2024; Islam et al., 2024); LLM-driven recommender systems and product knowledge graphs (Chen et al., 2024; Zhao et al., 2024); conversational commerce and customer service automation (Lim et al., 2025; Wirtz et al., 2023); machine translation for cross-border e-commerce (Gao et al., 2024); LLM-empowered review evaluation and summarization (Mathebula et al., 2024; Roumeliotis et al., 2024); and emerging agentic-AI architectures for

autonomous shopping and supply-chain decisions (Jin et al., 2024; Wang et al., 2024).

2.2 Adoption, behavior, and the consumer-side discourse

Parallel to the technical literature, a robust behavioral and adoption literature has examined how consumers and firms appropriate generative-AI tools. The Technology Acceptance Model (Davis, 1989), the Unified Theory of Acceptance and Use of Technology (Venkatesh et al., 2003), the Theory of Planned Behavior (Ajzen, 1991), and the Stimulus-Organism-Response (S-O-R) framework continue to anchor consumer-side studies of GenAI adoption (Foroughi et al., 2025; Mogaji et al., 2024; Yuan et al., 2025). However, recent contributions challenge whether classical adoption models remain sufficient for generative agents that exhibit anthropomorphic, creative, and even emotional behaviors. Mogaji et al. (2024) provocatively ask whether the technology acceptance model has reached its end in the era of generative AI; Aw et al. (2025) and Cham et al. (2025) extend behavioral models with anthropomorphism, creepiness, and trust constructs. Dwivedi et al. (2023) consolidate multidisciplinary perspectives on the opportunities, challenges, and policy

implications of generative conversational AI for research and practice.

2.3 Prior bibliometric reviews and what they leave unmapped

Several bibliometric reviews have addressed AI-related topics adjacent to ours, and Table 1 positions our work against the most relevant precedents.

Vlačić et al. (2021) synthesized 164 articles on AI in marketing, but their data window closed before the public emergence of ChatGPT and the broader generative paradigm. Mariani et al. (2022) conducted a systematic review of AI in marketing, consumer research, and psychology with similar pre-2022 boundaries. Kar et al. (2023) reviewed scientific and grey literature on generative AI in industrial applications using a structured topic-modeling approach but with limited bibliometric depth and no retail/e-commerce focus. Kumar et al. (2025) recently published a comprehensive bibliometric mapping of AI chatbot research in marketing covering 242 articles across 1996–2023 in the *Journal of Asia Business Studies*, providing a useful comparator for our methodology but explicitly bounding the corpus to "chatbot" terms and excluding generative use cases such as content generation, RAG, and agentic AI.

Table 1. Comparison with prior bibliometric reviews

Study	Topic	Window	N	Tool	Coverage of GenAI in retail/e-commerce
Vlačić et al. (2021)	AI in marketing	1990–2020	164	Bibliometrix	None (pre-GenAI)
Mariani et al. (2022)	AI in marketing/CR/Psy	–2021	~250	SLR	Marginal
Kar et al. (2023)	GenAI industrial applications	–2023	~150	Topic modeling	Limited; not retail-specific

Kumar et al. (2025)	AI chatbot in marketing	1996–2023	242	VOSviewer	Chatbot only; pre-GenAI majority
This study	GenAI in retail/e-commerce	2023–2026	445	VOSviewer + 7 lenses	Comprehensive, post-ChatGPT

2.4 Theoretical lenses for interpretation

To interpret the intellectual structure surfaced by the bibliometric analysis, we draw on three theoretical lenses. Bibliometric coupling theory (Kessler, 1963) and co-citation theory (Small, 1973) provide the foundational logic for treating shared references as evidence of shared intellectual community. The diffusion-of-innovation lens (Rogers, 2003) helps interpret the steep adoption curve and the geographic concentration of early adopters. Resource-based and dynamic-capability theories (Teece, 2007) help interpret the firm-level practitioner implications, particularly for retailers seeking to

build organizational capability around RAG, agentic AI, and conversational commerce.

3. Methodology

3.1 Database, search strategy, and corpus construction

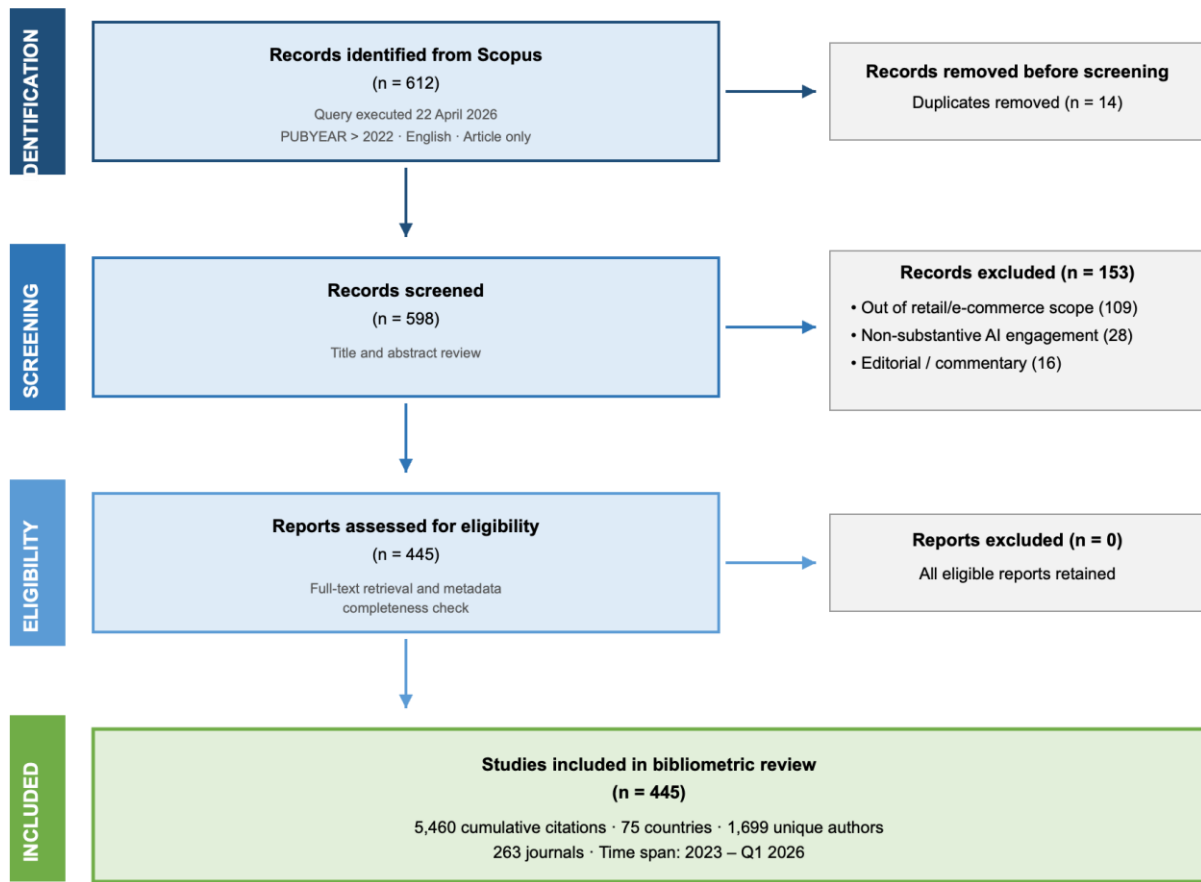
The Scopus database was selected as the primary information source given its broad multidisciplinary coverage, peer-review filter, and superior metadata quality for citation analyses (Mongeon & Paul-Hus, 2016). On 22 April 2026, the following Boolean query was executed on title, abstract, and keywords:

TITLE-ABS-KEY(("generative AI" OR "generative artificial intelligence" OR "large language model*" OR "LLM" OR "ChatGPT" OR "GPT-3" OR "GPT-4" OR "BERT" OR "transformer*" OR "retrieval augmented generation" OR "agentic AI") AND ("retail*" OR "e-commerce" OR "electronic commerce" OR "online shopping" OR "consumer" OR "customer experience" OR "recommender system*")) AND PUBYEAR > 2022 AND DOCTYPE(ar) AND LANGUAGE(English)

The query returned 612 records. Following the PRISMA 2020 framework (Page et al., 2021), records were screened in three stages (Figure 1): (i) duplicate removal (n = 14); (ii) title-and-abstract relevance screening to exclude papers that did not substantively engage with retail or e-commerce contexts (n = 153 excluded); and (iii)

full-text eligibility check to confirm methodological transparency and minimum metadata completeness (n = 0 excluded at this stage). The final corpus comprises 445 articles with 5,460 cumulative citations as of 22 April 2026.

Figure 1. PRISMA 2020 flow diagram of corpus construction



Adapted from Page et al. (2021). Coverage: 91.1% ref-source · 97.6% ref-author · 99.6% paper-country

Figure 1. PRISMA flow diagram.

3.2 Bibliometric procedures and software

Seven complementary bibliometric analyses were performed using VOSviewer 1.6.20 (van Eck & Waltman, 2010) and Pajek-formatted network files generated programmatically from the Scopus CSV export. Following Donthu et al.

(2021), full counting and association-strength normalization were used throughout to maintain comparability across analyses. Modularity-based clustering with default resolution (1.0) was applied to all networks. Table 2 reports the parameters and outcomes of each analysis.

Table 2. Bibliometric analyses, parameters, and outcomes

#	Analysis	Threshold	Items kept	Edges	Clusters
1	Keyword co-occurrence (author keywords + thesaurus)	min 3 occurrences	65	320	15 (5 macro + 10 micro)

2A	Co-citation of sources	min 30 citations	37	654	4
2B	Co-citation of first authors	min 20 citations	63	1,742	2
3A	Bibliographic coupling of documents	min 5 cited-by	170	3,077	9
3B	Bibliographic coupling of journals	min 2 documents	72	1,585	5
4A	Co-authorship of authors (Scopus IDs)	min 2 docs + 5 cites	39	114	11
4B	Co-authorship of countries	min 5 docs + 20 cites	31	172	6

3.3 Thesaurus harmonization and identifier disambiguation

To prevent fragmentation of the keyword network across morphological variants, a curated thesaurus of 26 mappings was applied prior to keyword co-occurrence (e.g., "large language model", "llm", "llms", "large language models (llms)", and "large language model (llm)" were unified to "large language models"; "generative artificial intelligence", "genai", and "gen-ai" were unified to "generative ai"). Co-authorship of authors used Scopus Author IDs to avoid name conflation, while co-citation of authors followed the VOSviewer convention of using the first author of each cited reference; we acknowledge this latter design conflates common short surnames such as "wang y" and "liu y", and we triangulate with full-name co-authorship to mitigate interpretation error.

3.4 Reference parsing and quality controls

Reference strings in Scopus exports follow a heuristic delimiter pattern that requires post-processing for source extraction. Our parsing pipeline resolved 24,288 of 26,653 (91.1%) reference strings to a journal source and 26,006 of 26,653 (97.6%) to a first author. Country

extraction from the Scopus *Affiliations* field resolved 443 of 445 (99.6%) papers to at least one country, with 75 distinct countries detected. These coverage rates are consistent with established bibliometric pipelines (Aria & Cuccurullo, 2017) and are conservative for aggregated network analysis.

4. Results

4.1 Annual publication trajectory and dataset overview

The 445-article corpus exhibits a near-doubling annual growth pattern (Figure 2). In 2023 (the first full year following the public release of ChatGPT), 12 articles were indexed. The annual count rose to 90 in 2024 and 218 in 2025, with 125 additional articles in the first quarter of 2026 alone — placing 2026 on track to exceed 2025 by year-end. Total cumulative citations stand at 5,460, with the average article in the corpus accumulating 12.27 citations and the median article 4 citations, indicative of the right-skew typical of bibliometric distributions. The 445 articles were authored by 1,699 unique authors at 1,012 institutions across 75 countries, signaling rapid global engagement.

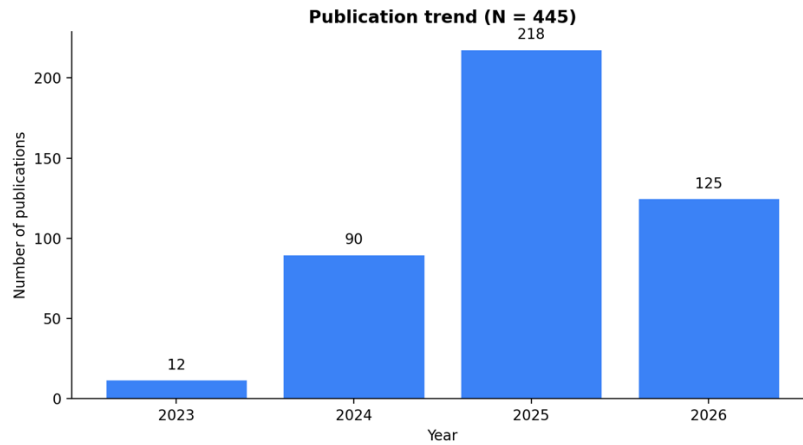


Figure 2. Annual publication trajectory, 2023–2026

4.2 Productive journals: the publishing infrastructure of the field

The 445 articles are distributed across 263 distinct journals, but a hub-periphery structure is apparent: 72 journals (27.4%) account for at least two articles each, and the top 10 outlets together publish 22.5% of the corpus (Table 3). The *Journal of Retailing and Consumer Services* is the single most productive outlet (24 articles, 538 citations within corpus), reflecting its established

positioning at the intersection of retail and consumer technology. *IEEE Access* (18 articles) and *Scientific Reports* (13 articles) provide rapid open-access publication channels for technically oriented contributions. The *Journal of Theoretical and Applied Electronic Commerce Research* (9), *Journal of Business Research* (7), *Acta Psychologica* (7), and *International Journal of Production Economics* (7) round out the productive tier and indicate the multidisciplinary character of the field.

Table 3. Top 10 most productive journals in the corpus

Rank	Journal	N papers	Within-corpus citations	TLS (bibcoupling)
1	Journal of Retailing and Consumer Services	24	538	1,251
2	IEEE Access	18	144	640
3	Scientific Reports	13	123	943
4	Journal of Theoretical and Applied Electronic Commerce Research	9	29	388
5	Journal of Business Research	7	155	620
6	Acta Psychologica	7	31	462

7	International Journal of Production Economics	7	12	290
8	Journal of Medical Internet Research	7	68	152
9	Applied Sciences (Switzerland)	6	25	342
10	Applied Marketing Analytics	6	14	39

Bibliographic coupling among these 72 journals reveals five thematic clusters — a marketing/consumer-services cluster (centered on JRCS), a technical/computer-science cluster (IEEE Access, Scientific Reports, Expert Systems with Applications), a hospitality/service cluster (Service Industries Journal, Tourism Management), an information-management cluster (IJIM, IJIM Data Insights), and a methods/operations cluster (IJ Production Economics, Knowledge-Based Systems). The hub-and-spoke topology indicates a still-consolidating publishing infrastructure; no single journal dominates the conversation, in contrast to more mature subfields where one or two flagship outlets typically anchor the discourse.

4.3 Co-citation of sources: the four theoretical pillars

Co-citation of cited sources, retaining the 37 sources cited at least 30 times within the corpus' references, reveals four theoretical pillars (Table

4). The dominant cluster comprises mainstream marketing and consumer-research journals: *Journal of Business Research* (160 within-corpus co-citations, TLS 1,285), *Journal of Retailing and Consumer Services* (107, TLS 933), *Computers in Human Behavior* (98, TLS 868), *Journal of Marketing Research* (84, TLS 827), *Journal of the Academy of Marketing Science* (82, TLS 801), and *Journal of Marketing* (75, TLS 729). A second cluster anchors the technical-CS heritage of the field around *IEEE Access* (106, TLS 465), *arXiv* (53, TLS 251), *Advances in Neural Information Processing Systems* (35, TLS 169), and *Expert Systems with Applications* (70, TLS 394). A third cluster captures the hospitality, services, and tech-society discourse, with *International Journal of Hospitality Management*, *Tourism Management*, *Technology in Society*, and *Telematics and Informatics*. The fourth cluster reflects the management-information-systems pillar with *MIS Quarterly* (80, TLS 724), *Information Systems Research* (52, TLS 482), *Decision Support Systems* (59, TLS 517), and *Management Science* (60, TLS 462).

Table 4. Top 12 co-cited sources (theoretical pillars)

Rank	Source	Citations	TLS	Pillar
1	Journal of Business Research	160	1,285	Marketing/CB
2	Journal of Retailing and Consumer Services	107	933	Marketing/CB
3	IEEE Access	106	465	Tech/CS

4	International Journal of Information Management	99	754	MIS/IS
5	Computers in Human Behavior	98	868	Marketing/CB
6	Journal of Marketing Research	84	827	Marketing/CB
7	Journal of the Academy of Marketing Science	82	801	Marketing/CB
8	MIS Quarterly	80	724	MIS/IS
9	Journal of Marketing	75	729	Marketing/CB
10	Sustainability	74	446	Tech/CS-OA
11	Expert Systems with Applications	70	394	Tech/CS
12	Technological Forecasting and Social Change	64	586	Tech-Society

The presence of four theoretically coherent yet structurally separate clusters provides the first quantitative evidence that GenAI-in-retail/e-commerce research is being constructed at the intersection of four distinct scholarly traditions, none of which historically read each other's flagship outlets at scale. This four-pillar architecture has direct implications for reviewer panel composition and journal positioning strategy (Section 6).

4.4 Co-citation of authors: the dual intellectual structure

The co-citation of first authors (63 authors with at least 20 citations within corpus references) reveals the central empirical finding of this study: the field is structured around two parallel intellectual traditions with limited cross-citation between them (Table 5).

The technical-AI tradition (Cluster A) is anchored by the foundational papers of the modern transformer era. Vaswani A. (51 co-citations, TLS 361) authored the seminal "Attention is All You Need" (Vaswani et al., 2017) introducing the transformer architecture. Devlin J. (50, TLS 353) authored BERT (Devlin et al., 2019). Brown T. (34, TLS 255) authored GPT-3 (Brown et al., 2020). Radford A. (34, TLS 273) authored GPT-1/2 and CLIP (Radford

et al., 2019). Touvron H. (27, TLS 232) authored LLaMA (Touvron et al., 2023). Wei J. (37, TLS 267) is associated with chain-of-thought prompting and instruction tuning. This cluster also includes high-frequency Chinese-surname first authors (Wang Y. 66, Liu Y. 61, Li Y. 57, Zhang Y. 57) whose conflated identity nonetheless tracks the volume of technical contributions emerging from China-based teams.

The adoption-behavioral tradition (Cluster B) is anchored by the methodological and theoretical workhorses of the behavioral marketing and information-systems literatures. Davis F.D. (23, TLS 150) is the originator of TAM (Davis, 1989). Venkatesh V. (38, TLS 246) led the UTAUT formulation (Venkatesh et al., 2003). Ajzen I. (24, TLS 165) authored the Theory of Planned Behavior (Ajzen, 1991). Hair J.F. (41, TLS 294) and Henseler J. (28, TLS 211) anchor the PLS-SEM methodological tradition that dominates empirical consumer-side studies. Dwivedi Y.K. (56, TLS 298) is the most-cited individual researcher in the corpus, having authored or co-authored multiple agenda-setting pieces on generative AI in IS and marketing (Dwivedi et al., 2023; Mogaji et al., 2024). Mogaji E. (25, TLS 140), Kshetri N. (29, TLS 161), and Paul J. (26, TLS 146) round out a behavioral cluster that is markedly more interconnected with marketing and consumer-behavior outlets than with technical-CS outlets.

Fornell C. (34, TLS 243) and Podsakoff P.M. (23, TLS 174) appear as methodological pillars

(variance-extracted reliability; common-method bias).

Table 5. Top 15 co-cited first authors and their intellectual cluster

Rank	First author	Citations	TLS	Cluster	Representative work
1	Wang Y.	66	459	A (Technical)	Various NLP/LLM
2	Liu Y.	61	380	A (Technical)	Various NLP/LLM
3	Li Y.	57	423	A (Technical)	Various NLP/LLM
4	Zhang Y.	57	385	A (Technical)	Various NLP/LLM
5	Dwivedi Y.K.	56	298	B (Behavioral)	GenAI policy & agenda
6	Vaswani A.	51	361	A (Technical)	Transformer (2017)
7	Devlin J.	50	353	A (Technical)	BERT (2019)
8	Hair J.F.	41	294	B (Behavioral)	PLS-SEM
9	Venkatesh V.	38	246	B (Behavioral)	UTAUT (2003)
10	Wei J.	37	267	A (Technical)	Chain-of-thought
11	Brown T.	34	255	A (Technical)	GPT-3 (2020)
12	Radford A.	34	273	A (Technical)	GPT-1/2; CLIP
13	Fornell C.	34	243	B (Behavioral)	Construct reliability
14	Kshetri N.	29	161	B (Behavioral)	GenAI policy
15	Henseler J.	28	211	B (Behavioral)	PLS-SEM

Modularity-based clustering returns exactly two clusters with default resolution — a compact, decisive partition that is rarely seen in modern bibliometric analyses (Donthu et al., 2021) and constitutes empirical confirmation of the long-suspected "two communities" structure. Importantly, the within-cluster TLS density is nearly an order of magnitude higher than the between-cluster TLS density: behavioral authors

disproportionately co-cite Davis, Venkatesh, Hair, and Henseler, while technical authors disproportionately co-cite Vaswani, Devlin, Brown, and Touvron. Cross-tradition citations exist but are concentrated in a small number of bridging actors — most notably Dwivedi Y.K., Mogaji E., and Kshetri N., whose work explicitly synthesizes technical and behavioral perspectives on generative AI.

4.5 Bibliographic coupling of documents: thematic frontiers and high-impact works

The bibliographic coupling network of documents (170 articles with at least 5 citations each, 3,077 edges) returns nine thematic clusters that map onto distinct application and methodological domains (Table 6 lists the highest-impact paper in each cluster). The largest cluster centers on consumer behavior and adoption of GenAI in marketing, anchored by Hermann and Puntoni's (2024) *Journal of Business Research* piece (128 citations) and Mogaji et al.'s (2024) *International Journal of Contemporary Hospitality Management* challenge to TAM (130 citations). A second cluster captures content generation and brand authenticity, with Brüns and Meißner (2024) and

Islam et al. (2024). A third cluster focuses on LLMs and recommender systems (Zhao et al., 2024, 266 citations) including LLMs-on-graphs frontiers (Jin et al., 2024, 92 citations). A fourth captures AI-driven sentiment analysis and review evaluation (Roumeliotis et al., 2024, 92 citations). A fifth concentrates on operations and supply-chain applications (Fosso Wamba et al., 2024). A sixth emphasizes service automation, chatbots, and conversational commerce (Wirtz et al., 2023; Lim et al., 2025). A seventh captures innovation and strategy (wael Al-Khatib, 2023, 154 citations; Cillo & Rubera, 2025). An eighth captures privacy, ethics, and policy (Ali et al., 2025; Jaidka et al., 2025). The ninth and most recent cluster — with average publication year 2025.6 — captures the agentic-AI and RAG frontier.

Table 6. Highest-impact paper per bibcoupling cluster

Cluster	Theme	Anchor paper	Citations
1	Consumer behavior & adoption	Mogaji et al. (2024) IJCHM	130
2	Content generation & authenticity	Brüns & Meißner (2024) JRCS	—
3	LLMs & recommender systems	Zhao et al. (2024) IEEE TKDE	266
4	NLP/sentiment/review eval.	Roumeliotis et al. (2024) NLP Journal	92
5	Operations & supply chain	Fosso Wamba et al. (2024) IJ Production Research	—
6	Service & conversational commerce	Wirtz et al. (2023) Italian J. Marketing	34
7	Innovation & strategy	Al-Khatib (2023) Technology in Society	154
8	Privacy, ethics, policy	Ali et al. (2025) ACM Computing Surveys	31
9	Agentic AI / RAG frontier	Jin et al. (2024) IEEE TKDE (LLMs on Graphs)	92

Notably, Hermann and Puntoni's (2024) "Artificial intelligence and consumer behavior: From predictive to generative AI" emerges as the single highest-bibcoupling-strength article in the corpus, a position consistent with its agenda-setting framing of the predictive-to-generative

shift. Foroughi et al. (2026), despite having only 5 citations at the time of analysis, exhibits the second-highest bibcoupling strength — a signature pattern of an emerging high-impact paper that is being heavily cited by the most

recent 2026 articles but has not yet accumulated raw citation volume.

4.6 Keyword co-occurrence: thematic clusters and the emerging-concept frontier

The author-keyword co-occurrence network, after thesaurus harmonization, retains 65 keywords (out of 1,606 raw unique keywords) at the minimum-three-occurrence threshold (Table 7). The five macro-clusters that emerge from modularity-based clustering (resolution 1.0, returning 15 clusters of which 5 are theoretically substantial and 10 are micro-clusters of 2–4 keywords each) are summarized below.

Cluster 1 — GenAI core and customer outcomes (anchor: *generative ai*, occ = 115, TLS = 152). Keywords: *trust, customer experience, customer satisfaction, online reviews, purchase intention, consumer behavior*. This cluster captures the consumer-side empirical core of the field.

Cluster 2 — AI chatbots and adoption (anchor: *artificial intelligence*, occ = 73; *chatgpt*, occ = 56). Keywords: *chatbots, technology adoption,*

social media, anthropomorphism. Captures the chatbot-as-conversational-agent inheritance that bridges to the Kumar et al. (2025) literature.

Cluster 3 — NLP/ML methods (anchor: *natural language processing*, occ = 33). Keywords: *sentiment analysis, machine learning, deep learning, transformers, bert, fine-tuning, contrastive learning*. Captures the technical-methods backbone of empirical work.

Cluster 4 — E-commerce and customer experience (anchor: *e-commerce*, occ = 37, average year 2024.97). Keywords: *customer engagement, personalization, e-commerce platform, digital marketing*. Average publication year 2025.4 makes this the youngest of the five macro-clusters.

Cluster 5 — LLMs and recommendation (anchor: *large language models*, occ = 144, the largest single keyword in the corpus). Keywords: *recommender systems, prompt engineering, reinforcement learning, multi-agent systems*. The high TLS (194) and centrality of "large language models" makes this the structural backbone of the network.

Table 7. Top 25 keywords by occurrence (after thesaurus)

Rank	Keyword	Occ.	TLS	Avg. year	Cluster
1	large language models	144	194	2025.08	5
2	generative ai	115	152	2024.99	1
3	artificial intelligence	73	144	2024.71	2
4	chatgpt	56	103	2024.43	2
5	e-commerce	37	82	2024.97	4
6	natural language processing	33	90	2024.79	3
7	sentiment analysis	25	62	2024.72	3
8	machine learning	23	59	2024.83	3
9	chatbots	22	48	2024.77	2
10	deep learning	19	59	2024.84	3
11	customer experience	16	44	2025.06	1
12	recommender systems	15	32	2024.93	5
13	trust	15	26	2025.27	1

14	consumer behavior	14	22	2025.14	1
15	digital marketing	9	23	2024.56	4
16	transformers	8	25	2025.00	3
17	customer engagement	7	8	2025.43	4
18	technology adoption	7	17	2025.00	2
19	prompt engineering	6	11	2024.83	5
20	online reviews	6	6	2025.33	1
21	personalization	5	15	2025.20	4
22	reinforcement learning	5	7	2025.00	5
23	fake review detection	4	9	2025.50	3
24	fine-tuning	4	10	2025.00	3
25	game theory	4	6	2025.75	—

The emerging-concept frontier: agentic AI, RAG, and conversational AI

Beyond the five macro-clusters, ten micro-clusters of 2–4 keywords each surface emerging

concepts whose average publication year exceeds 2025.3 — the empirical "future frontier" of the field. Table 8 catalogs these emerging concepts.

Table 8. Emerging concepts on the research frontier (average year > 2025.3)

Emerging concept	Occ.	TLS	Avg. year	Interpretation
agentic ai	3	5	2025.67	Autonomous LLM agents for shopping/supply chain
retrieval-augmented generation	3	7	2025.67	Grounded LLM responses with citation
conversational ai	3	3	2025.67	Beyond first-gen chatbots
e-commerce platform	3	2	2025.67	Platform-level GenAI integration
fake review detection	4	9	2025.50	Defensive applications of LLMs
customer engagement	7	8	2025.43	Engagement reframed for GenAI era
multi-agent systems	3	5	2025.33	Coordinated agentic architectures

schema theory	3	3	2025.33	Cognitive lens for AI design output
graph neural networks	3	5	2025.00	LLM-on-graphs convergence
game theory	4	6	2025.75	Strategic interaction models

These ten emerging concepts cluster around four tightly coupled architectural paradigms.

Agentic AI — the deployment of LLM-driven agents capable of autonomous goal-directed action in retail and e-commerce environments — appears as both a stand-alone keyword (occ 3, average year 2025.67) and indirectly through "multi-agent systems" (occ 3, average year 2025.33). Wang et al. (2024) demonstrate an LLM-empowered "smart customer service" agent in unmanned retail stores; Geng and Chen (2025) deploy LLMs with genetic algorithms for vehicle–drone cooperative delivery scheduling; Sun et al. (2025) propose LLM-as-promoter agents for e-commerce content generation. The agentic-AI frontier is structurally coupled to RAG and graph-based knowledge representation, suggesting that the next 12–24 months will see increasing convergence among these architecturally adjacent capabilities.

Retrieval-augmented generation (RAG) appears under two thesaurus-equivalent labels (*retrieval augmented generation*, average year 2024.33; *retrieval-augmented generation*, average year 2025.67) with combined TLS of 14, indicating an early but rapidly accelerating empirical interest. RAG is structurally important for retail/e-commerce because it addresses the persistent hallucination problem of pure LLM-based product answers by grounding responses in verifiable product catalogs and customer-specific knowledge bases (Ali et al., 2025; Chen et al., 2024). The bridge between RAG and graph-neural-networks (occ 3, average year 2025.00) is documented in Jin et al. (2024) and represents a high-leverage intersection where retrieval precision, knowledge-graph reasoning, and generative fluency converge.

Conversational AI (occ 3, average year 2025.67) marks an explicit lexical pivot away from "chatbots" (occ 22, average year 2024.77) toward more architecturally and pragmatically sophisticated systems capable of multi-turn reasoning, emotional sensitivity (Luo et al., 2025), and service-recovery management (Jeon & Kim, 2025). The terminological transition is itself diagnostic: the field is leaving the "chatbot" era and entering a more architecturally heterogeneous conversational-AI era.

Fake-review detection (occ 4, TLS 9, average year 2025.50) emerges as a defensive sub-frontier: as GenAI is increasingly used to generate fraudulent product reviews, parallel research is using LLMs to detect them (Massenon et al., 2025; Roumeliotis et al., 2024). This adversarial dynamic represents a research frontier of considerable practical importance to retailers managing review-platform integrity.

The overlay temporal visualization of the keyword network (corpus average year approximately 2024.9, with overlay range 2023.0–2026.0) shows a clearly identifiable yellow-to-green gradient in which 2026-published articles cluster preferentially around agentic AI, conversational AI, RAG, customer engagement, schema theory, and multi-agent systems — visually corroborating the emerging-frontier interpretation.

4.7 Geographic distribution and country impact differentials

The 445 articles were authored from 75 countries with at least one contribution. Table 9 reports the 18 countries that contributed at least 7 articles and at least 20 cumulative citations.

Table 9. Country productivity and impact (top 20; sorted by citations per paper)

Rank by impact	Country	Papers	Citations	Cites/paper	MCP	SCP	MCP ratio	TLS
1	Romania	6	536	89.33	2	4	0.33	6
2	United Arab Emirates	9	795	88.33	9	0	1.00	24
3	Taiwan	10	616	61.60	8	2	0.80	36
4	France	14	764	54.57	12	2	0.86	27
5	Malaysia	17	829	48.76	14	3	0.82	48
6	Brazil	6	253	42.17	3	3	0.50	4
7	Hong Kong	7	284	40.57	5	2	0.71	6
8	Germany	19	676	35.58	10	9	0.53	23
9	United Kingdom	36	1,062	29.50	30	6	0.83	65
10	India	51	1,491	29.24	26	25	0.51	70
11	Singapore	11	275	25.00	9	2	0.82	25
12	Turkey	11	267	24.27	5	6	0.45	6
13	Pakistan	8	186	23.25	8	0	1.00	12
14	Sweden	6	129	21.50	5	1	0.83	14
15	Macau	5	94	18.80	5	0	1.00	23
16	Greece	7	130	18.57	1	6	0.14	10
17	Italy	16	228	14.25	10	6	0.63	18
18	United States	100	1,296	12.96	40	60	0.40	76
19	Australia	20	187	9.35	10	10	0.50	36
20	Canada	9	90	10.00	4	5	0.44	13

Notes: MCP = multiple-country publications; SCP = single-country publications; TLS = total link strength in country co-authorship network. China and the United States not in top-impact rank but reported separately given their volume leadership.

Three country-level impact patterns

Three distinct country-level patterns emerge

Pattern 1 — Volume-led portfolios: China (112 papers, 7.48 cites/paper) and South Korea (35 papers, 7.51) lead in raw output but exhibit per-paper citation impact below the corpus mean. Their relatively low international collaboration ratios (China MCP = 0.39; South Korea MCP = 0.43) suggest that volume is being generated

primarily through domestic networks. The United States (100 papers, 12.96 cites/paper) presents an intermediate pattern: high volume combined with the highest absolute total link strength (TLS = 76) but a moderate MCP ratio (0.40), reflecting a domestically large but globally active research base.

Pattern 2 — Impact-led portfolios: India (51 papers, 29.24 cites/paper) and the United Kingdom (36 papers, 29.50) demonstrate that

high productivity and high impact can coexist when international collaboration is intense. The United Kingdom's MCP ratio of 0.83 — meaning that 83% of UK-affiliated papers in the corpus involve at least one non-UK co-author — is the highest among countries with at least 30 papers and is an order of magnitude above the corpus median. India combines a 0.51 MCP ratio with a sizable absolute TLS (70), indicating that Indian researchers are increasingly central nodes in the global GenAI-retail/e-commerce collaboration graph. The presence of Mogaji, Dwivedi, Kar, and other UK- and India-affiliated authors among the most-cited individuals in the corpus reinforces this pattern.

Pattern 3 — "Small-but-impactful" portfolios: The United Arab Emirates (9 papers, 88.33 cites/paper, MCP ratio 1.0), Romania (6 papers, 89.33 cites/paper), Taiwan (10 papers, 61.60), France (14 papers, 54.57), and Malaysia (17 papers, 48.76) demonstrate a striking pattern in which a small number of high-impact contributions, frequently authored in international collaboration, generate disproportionate citation impact. The UAE's 100% MCP ratio (every UAE-affiliated paper

involves international co-authorship) is consistent with the country's strategy of recruiting senior international researchers. Romania's high impact is driven by a small number of highly-cited contributions to the GenAI-strategy and innovation literature. Malaysia's strong showing reflects an active research community around behavioral GenAI adoption (Foroughi, Ghobakhloo, Wen, and colleagues).

The country co-authorship network (Figure 3) reveals six clusters: an Anglo-American cluster (US–UK–Australia–Canada), a South-Asian cluster (India–UAE–Saudi Arabia–Pakistan), an East-Asian cluster (China–Hong Kong–Macau–Taiwan–South Korea), a Continental-European cluster (Germany–France–Italy–Spain–Sweden–Greece–Poland), a Southeast-Asian cluster (Malaysia–Singapore–Thailand–Vietnam–Indonesia), and a Middle-East/MENA cluster (UAE–Jordan–Morocco–Turkey). Each cluster has emerging gateway countries: Singapore, Malaysia, and the UAE function as inter-cluster bridges with elevated TLS relative to volume.

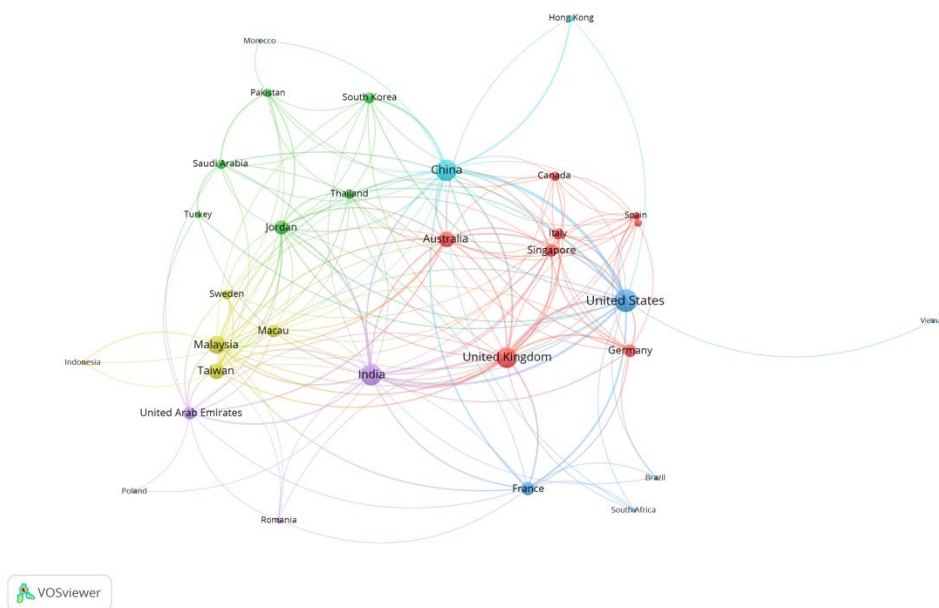


Figure 3. Country co-authorship network

4.8 Author co-authorship: a fragmented but emerging collaboration architecture

The co-authorship network of authors (39 authors with at least 2 papers and 5 citations each, 11 clusters, 114 edges) presents a fragmented architecture characteristic of a young field still organizing around productive teams. Yogesh K. Dwivedi (Swansea University, UK) is the most-published author in the corpus (7 papers, 702 within-corpus citations) and the most-co-cited researcher overall. Emmanuel Mogaji (Keele University, UK; previously Greenwich), Keng-Boon Ooi (UCSI University, Malaysia), Garry Wei-Han Tan (UCSI), Jochen Wirtz (NUS Singapore), Mohammed A. Al-Sharafi (Universiti Tenaga Nasional, Malaysia), and Behzad Foroughi anchor the most productive teams. The Wong–Raman cluster (Sunway University) also emerges as a productive Malaysian node. Notably, none of the 11 clusters exceeds 6 authors, and several contain only 2–3 mutually connected authors, signaling a still-fragmented collaboration space with substantial room for consolidation as the field matures.

5. Discussion

The bibliometric analysis presented above produces five interpretive findings whose joint significance reframes how the GenAI-in-retail/e-commerce research conversation should be understood and curated.

(1) The field is structurally bifurcated. The two-cluster co-citation structure of authors — with negligible cross-cluster citation density — constitutes the first quantitative demonstration that GenAI scholarship in retail/e-commerce is being produced by two effectively independent intellectual communities. The technical-AI tradition (Vaswani, Devlin, Brown, Radford, Touvron, Wei) provides the architectural foundations and the empirical NLP backbone; the adoption-behavioral tradition (Davis, Venkatesh, Ajzen, Hair, Henseler, Dwivedi, Mogaji) provides the consumer-side theoretical apparatus and the empirical methodology. This bifurcation is structurally similar to the well-documented split in earlier waves of e-commerce

research between the technical-IS and consumer-marketing traditions (Pavlou, 2003) but is more pronounced because the technical lineage is rooted in deep-learning publications that were never read at scale by marketing scholars before 2022. The implication is twofold: editors and reviewers should consciously construct cross-tradition review panels for GenAI papers, and authors should consciously cite across traditions to accelerate community integration. The bridging function performed by Dwivedi, Mogaji, and Kshetri provides a template that other authors can emulate.

(2) The frontier is agentic. Average-year analysis of keywords and bibcoupling clusters consistently positions agentic AI, RAG, conversational AI, and multi-agent systems on the leading edge of the field's lexical evolution. The structural coupling between RAG and graph-neural-networks (Jin et al., 2024) and between agentic AI and operations/supply-chain applications (Geng & Chen, 2025; Wang et al., 2024) suggests that the next wave of high-impact contributions will focus on architecturally integrated systems rather than on stand-alone LLM applications. From a research-strategy perspective, papers framed around agentic-AI deployment in retail (autonomous shopping assistants, dynamic-pricing agents, supply-chain co-pilots) and around grounded-RAG architectures for product answer-generation are likely to occupy a high-leverage position in the 2026–2028 publication cycle.

(3) Country impact is highly heterogeneous, and small countries can punch far above their weight. The ratio between the highest and lowest citations-per-paper in our sample exceeds 35:1 (Romania 89.3 vs. Indonesia 2.5), and the impact-leaders are not the volume-leaders. The "small-but-impactful" pattern observed in the UAE, France, Taiwan, Malaysia, and Romania has direct implications for research-policy: high-impact GenAI scholarship can be cultivated through targeted international-collaboration strategies and selective senior-researcher recruitment, not only through volume expansion. Conversely, the China-South Korea pattern of high-volume domestic publication with below-mean per-paper impact suggests that volume

strategies, in the absence of intensive international collaboration, do not by themselves translate to scholarly impact in this nascent field. India's combination of high volume and high impact (51 papers at 29.2 cites/paper) demonstrates that the two strategies are not mutually exclusive.

(4) The chatbot literature is being absorbed and renamed. The lexical pivot from "chatbots" (average year 2024.77) to "conversational AI" (average year 2025.67) and the embedding of chatbot scholarship within the broader GenAI cluster signal an important conceptual transition: the AI-chatbot research stream documented by Kumar et al. (2025) is being absorbed into a richer GenAI conversation in which conversational interfaces are one capability among many. Authors continuing to publish under the "chatbot" framing should consider repositioning their work within the broader conversational-AI or agentic-AI vocabulary to maintain visibility.

(5) The journal portfolio is multidisciplinary and unconsolidated. No single outlet dominates the corpus — the most productive journal (JRCS) accounts for only 5.4% of articles — and the four-pillar co-citation structure of cited sources confirms that contributors are drawing on disparate scholarly traditions. For authors targeting Q1 placement, this implies that successful papers are typically those that explicitly bridge traditions: combining a technically credible methods section with theoretically grounded behavioral framing or combining a robust empirical study with policy-relevant practical implications. JRCS, JBR, IJIM, IJIM Data Insights, and Technology in Society emerge as outlets where bridging papers have demonstrably succeeded in 2023–2026.

6. Implications

6.1 Theoretical implications

First, the documentation of the dual technical-behavioral intellectual structure provides a rigorous foundation for cross-tradition theory development. Rather than treating GenAI as either a technological artifact (technical tradition) or a consumer stimulus (behavioral

tradition), future theory should explicitly model the two-way interaction: how the architectural choices of a generative system (model size, fine-tuning, RAG grounding, agentic autonomy) condition the consumer responses traditionally captured by TAM/UTAUT/S-O-R, and how observed consumer responses feed back into design and architectural decisions. Hermann and Puntoni (2024), Mogaji et al. (2024), and Cillo and Rubera (2025) provide initial theoretical scaffolding; substantial conceptual work remains.

Second, the positioning of agentic AI and RAG as empirically grounded research frontiers calls for theoretical apparatus capable of describing autonomous, multi-step interactions between consumers and AI agents. Classical TAM/UTAUT models assume consumers as decision-makers evaluating a tool; agentic-AI deployment increasingly delegates substantive decision-making to the agent itself. Theoretical frameworks for "human-AI co-decision" (Vaccaro et al., 2024 in *Nature Human Behaviour*), "responsibility attribution" (Yue et al., 2023), and "schema theory" (Han et al., 2026) emerge as candidate theoretical bridges that warrant systematic theoretical development.

Third, the country-impact differentials documented above invite theoretical engagement with the geography of innovation and intellectual leadership in nascent fields. The "small-but-impactful" pattern is consistent with theoretical work on knowledge brokerage and gateway nodes in collaboration networks (Burt, 2004) but has not been documented in this empirical setting before.

Fourth, the four-pillar source structure (marketing/CB, technical/CS, hospitality/services, MIS/IS) suggests that GenAI in retail/e-commerce is most usefully theorized as an interdisciplinary crossroads rather than as an extension of any single parent tradition. Theoretical contributions that explicitly synthesize across pillars — for instance, integrating IS-flavored adoption theory with marketing-flavored experience theory and CS-flavored architectural detail — are likely to occupy high-leverage positions in the conversation.

6.2 Practical implications

For retail and e-commerce executives, the bibcoupling-cluster map provides a capability roadmap: well-developed clusters (LLMs/recommender systems, NLP/sentiment, content generation) indicate areas where the academic literature can directly inform deployment design, while emerging-frontier clusters (agentic AI, RAG, multi-agent orchestration) indicate areas where research is still active and where early-mover firms can generate competitive advantage by participating in research-practice partnerships. The strong showing of *IEEE TKDE* (Zhao et al., 2024; Jin et al., 2024) and *Expert Systems with Applications* on the technical side, and JRCS, JBR, and IJIM on the behavioral side, makes the "bridging" literature accessible to practitioners through both sets of channels.

For research administrators and funding bodies, the country-impact differential surfaces the high return on international-collaboration funding mechanisms. The UAE, France, Malaysia, and India patterns suggest that targeted internationalization, rather than volume expansion alone, is the more effective lever for building scholarly impact in a nascent field. National funding agencies in countries with growing but still-low-impact portfolios (Vietnam, Indonesia, Saudi Arabia, Poland) should consider co-funded international-collaboration programs as a primary instrument.

For academic researchers, the journal-portfolio map and the bridging-author template (Dwivedi, Mogaji, Kshetri) provide actionable strategic guidance. Researchers seeking to enter the field should target the bridging journals (JBR, IJIM, IJIM Data Insights, JRCS) with theoretically grounded studies that explicitly cite across the technical-behavioral divide. The fragmented co-authorship architecture (39 authors, 11 clusters) means that opportunities for productive collaboration with established teams remain abundant.

For platform operators and policymakers, the rapid emergence of fake-review detection, ethics, privacy, and policy concerns within the corpus signals that responsible-AI considerations are not a separate compliance

overhead but an integral part of the research conversation. Ali et al. (2025), Jaidka et al. (2025), and Kunz and Wirtz (2024) on corporate digital responsibility provide initial frameworks; integration into product and policy design should be early rather than retrofitted.

7. Conclusion, limitations, and future research

This study provides the first comprehensive bibliometric mapping of generative-AI scholarship at the retail and e-commerce intersection. Drawing on 445 Scopus-indexed articles published between 2023 and the first quarter of 2026 and applying seven complementary VOSviewer analyses, we documented the field's exponential growth, its dual technical-behavioral intellectual structure, its five thematic macro-clusters and ten micro-clusters of emerging concepts, its multidisciplinary publishing infrastructure, and its highly heterogeneous country-impact landscape. The agentic-AI/RAG/conversational-AI frontier emerges as the empirically grounded leading edge of the field, and the small-but-impactful country pattern reframes how research-strategy decisions should be considered in nascent fields.

The study has five limitations. First, coverage is restricted to the Scopus database; complementary searches in Web of Science, Dimensions, and Lens.org would extend the corpus and triangulate the findings, particularly for highly recent 2026 work that Scopus may index with delay. Second, the four-year window (2023–2026) is narrow; while this is appropriate given the post-ChatGPT focus, longitudinal continuity with pre-2023 AI scholarship requires additional analysis. Third, the first-author convention for co-citation analysis conflates common short surnames (most visibly "wang y", "liu y", "li y"), and although we triangulated with Scopus-ID-based co-authorship for author-level analyses, residual conflation remains in the co-citation findings. Fourth, English-language and peer-reviewed-article filters excluded conference papers, preprints, books, industry reports, and non-English work; given the rapid GenAI development cycle, preprint exclusion

may understate the true frontier. Fifth, reference-string parsing achieved 91.1% source resolution and 97.6% first-author resolution, leaving residual ~9% information loss that is acceptable for aggregate network analysis but bounds inference at the individual-edge level.

Five high-leverage future research directions follow directly from the analysis. (1) Theoretical integration: developing formal frameworks that integrate architectural design choices (RAG grounding, agentic autonomy, retrieval scope) with consumer-response constructs (trust, engagement, purchase intention). (2) Frontier consolidation: empirical studies of agentic AI and RAG in retail/e-commerce settings, particularly large-N field experiments and longitudinal deployment studies. (3) Adversarial dynamics: systematic study of the GenAI-fake-review and GenAI-fraud-detection arms race, including economic-equilibrium and policy-design considerations. (4) Cross-cultural and emerging-market work: consumer responses to GenAI-driven commerce vary substantially across cultural contexts; ASEAN, MENA, Latin-American, and African empirical work remains under-represented relative to Western and East-Asian work. (5) Welfare and ethical evaluation: beyond firm-level efficiency gains, systematic measurement of consumer-welfare impact, including AI-driven dark patterns, algorithmic price discrimination, and labor-displacement effects, deserves substantive empirical attention. These directions, taken together, define a research agenda capable of carrying the field through its next phase of consolidation.

References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Al-Khatib, A. W. (2023). Drivers of generative artificial intelligence to fostering exploitative and exploratory innovation: A TOE framework. *Technology in Society*, 75, 102403. <https://doi.org/10.1016/j.techsoc.2023.102403>
- Ali, W., Yang, M., Ali, X., Wang, S., & Wu, J. (2025). Privacy-preserved and responsible recommenders: From conventional defense to federated learning and blockchain. *ACM Computing Surveys*, 57(7), Article 174. <https://doi.org/10.1145/3708982>
- Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959–975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Aw, E. C.-X., Cham, T.-H., Tan, G. W.-H., & Ooi, K.-B. (2025). The age of generative AI conversational commerce. *Journal of Retailing and Consumer Services*, 78, 103746.
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
- Brüns, J. D., & Meißner, M. (2024). Do you create your content yourself? Using generative artificial intelligence for social media content creation diminishes perceived brand authenticity. *Journal of Retailing and Consumer Services*, 79, 103790. <https://doi.org/10.1016/j.jretconser.2024.103790>
- Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, 110(2), 349–399.
- Chen, J., Xu, Y., & Liu, T. (2024). Relation labeling in product knowledge graphs with large language models for e-commerce. *International Journal of Machine Learning and Cybernetics*, 15(11), 5275–5290. <https://doi.org/10.1007/s13042-024-02274-5>
- Cillo, P., & Rubera, G. (2025). Generative AI in innovation and marketing processes: A roadmap of research opportunities. *Journal of the Academy of Marketing Science*, 53(4), 685–704.

<https://doi.org/10.1007/s11747-024-01044-7>

- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT 2019*, 4171–4186.
- 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>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., et al. (2023). "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, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Foroughi, B., Nhan, P. V., Iranmanesh, M., Ghobakhloo, M., Nilashi, M., & Yadegaridehkordi, E. (2025). Determinants of intention to use ChatGPT for professional services: A PLS-ANN analysis. *Journal of Retailing and Consumer Services*. (Advance online publication).
- Fosso Wamba, S., Guthrie, C., Queiroz, M. M., & Minner, S. (2024). ChatGPT and generative artificial intelligence: An exploratory study of key benefits and challenges in operations and supply chain management. *International Journal of Production Research*, 62(16), 5676–5696. <https://doi.org/10.1080/00207543.2023.2294116>
- Gao, D., Hu, P., & Huang, Y. (2024). LLMs-based machine translation for E-commerce. *Expert Systems with Applications*, 256, 125087. <https://doi.org/10.1016/j.eswa.2024.125087>
- Geng, M., & Chen, A. (2025). Optimized scheduling for multi-drop vehicle–drone collaboration with delivery constraints using large language models and genetic algorithms with symmetry principles. *Symmetry*, 17(6), 934. <https://doi.org/10.3390/sym17060934>
- Guha, A., Grewal, D., Kopalle, P. K., Haenlein, M., Schneider, M. J., Jung, H., et al. (2021). How artificial intelligence will affect the future of retailing. *Journal of Retailing*, 97(1), 28–41.
- Han, W., Lim, A., & Kim, S. (2026). "Make an effort and show me the love!" Understanding consumer aversion to AI fashion design through the lens of schema theory. *Clothing and Textiles Research Journal*. (Advance online publication). <https://doi.org/10.1177/0887302X251347924>
- Hermann, E., & Puntoni, S. (2024). Artificial intelligence and consumer behavior: From predictive to generative AI. *Journal of Business Research*, 180, 114720. <https://doi.org/10.1016/j.jbusres.2024.114720>
- Islam, T., Miron, A., Liu, X., & Li, Y. (2024). Transforming digital marketing with generative AI. *Computers*, 13(7), 168. <https://doi.org/10.3390/computers13070168>
- Jaidka, K., Chen, Y., & Pang, A. (2025). Misinformation, disinformation, and generative AI: Implications for perception and policy. *Digital Government: Research and Practice*, 6(1). <https://doi.org/10.1145/3689372>
- Jeon, E., & Kim, J. (2025). When time is crucial: Effects of compensation type and anthropomorphic chatbots on chatbot

- service recovery satisfaction. *Journal of Research in Interactive Marketing*. (Advance online publication). <https://doi.org/10.1108/JRIM-05-2024-0238>
- Jin, B., Liu, G., Han, C., Jiang, M., Ji, H., & Han, J. (2024). Large language models on graphs: A comprehensive survey. *IEEE Transactions on Knowledge and Data Engineering*, 36(12), 8622–8642. <https://doi.org/10.1109/TKDE.2024.3469578>
- Kar, A. K., Varsha, P. S., & Rajan, S. (2023). Unravelling the impact of generative artificial intelligence (GAI) in industrial applications: A review of scientific and grey literature. *Global Journal of Flexible Systems Management*, 24, 659–689. <https://doi.org/10.1007/s40171-023-00356-x>
- Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 14(1), 10–25.
- Kshetri, N., Dwivedi, Y. K., Davenport, T. H., & Panteli, N. (2024). Generative artificial intelligence in marketing: Applications, opportunities, challenges, and research agenda. *International Journal of Information Management*, 75, 102716. <https://doi.org/10.1016/j.ijinfomgt.2023.102716>
- Kumar, S., Kumar, P., Priya, P., & Patel, G. (2025). Mapping the evolution of artificial intelligence (AI) chatbot in marketing: A bibliometric analysis. *Journal of Asia Business Studies*, 19(6), 1253–1275. <https://doi.org/10.1108/JABS-12-2024-0664>
- 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>
- Lim, J. S., Schneider, E., Grover, M., Zhang, J., & Peters, D. (2025). Effects of AI versus human source attribution on trust and forgiveness in the identical corporate apology statement for a data breach scandal. *Public Relations Review*, 51(1), 102520. <https://doi.org/10.1016/j.pubrev.2024.102520>
- Luo, X., Wang, Y., & Zhang, L. (2025). Seeking emotional and mental health support from generative AI: Mixed-methods study of ChatGPT user experiences. *JMIR Mental Health*, 12, e77951. <https://doi.org/10.2196/77951>
- Mariani, M., Vega, R. P., & Wirtz, J. (2022). AI in marketing, consumer research and psychology: A systematic literature review and research agenda. *Psychology & Marketing*, 39(4), 755–776. <https://doi.org/10.1002/mar.21619>
- Massenon, R., Almeida, J., & Kumar, S. (2025). "My AI is lying to me": User-reported LLM hallucinations in AI mobile apps reviews. *Scientific Reports*, 15, 15416. <https://doi.org/10.1038/s41598-025-15416-8>
- Mathebula, M., Mosola, N., Modupe, A., & Owolawi, P. A. (2024). ChatGPT as a text annotation tool to evaluate sentiment analysis on South African financial institutions. *IEEE Access*, 12, 144016–144030. <https://doi.org/10.1109/ACCESS.2024.3464374>
- Mogaji, E., Viglia, G., Srivastava, P., & Dwivedi, Y. K. (2024). Is it the end of the technology acceptance model in the era of generative artificial intelligence? *International Journal of Contemporary Hospitality Management*, 36(10), 3324–3343. <https://doi.org/10.1108/IJCHM-08-2023-1271>
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213–228.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020

- statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71.
- Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, 7(3), 101–134.
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI Blog*, 1(8).
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
- Roumeliotis, K. I., Tselikas, N. D., & Nasiopoulos, D. K. (2024). LLMs in e-commerce: A comparative analysis of GPT and LLaMA models in product review evaluation. *Natural Language Processing Journal*, 6, 100056. <https://doi.org/10.1016/j.nlp.2024.100056>
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4), 265–269.
- Sun, X., Shi, K., Tang, H., Wang, D., Xu, G., & Li, Q. (2025). Educating language models as promoters: Multi-aspect instruction alignment with self-augmentation. *IEEE Transactions on Knowledge and Data Engineering*, 37(8), 4564–4577. <https://doi.org/10.1109/TKDE.2025.3569585>
- Teece, D. J. (2007). Explicating dynamic capabilities. *Strategic Management Journal*, 28(13), 1319–1350.
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., et al. (2023). LLaMA: Open and efficient foundation language models. *arXiv preprint*, arXiv:2302.13971.
- Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478.
- Vlačić, B., Corbo, L., Costa e Silva, S., & Dabić, M. (2021). The evolving role of artificial intelligence in marketing: A review and research agenda. *Journal of Business Research*, 128, 187–203. <https://doi.org/10.1016/j.jbusres.2021.01.055>
- Wang, W., Li, X., & Zhang, K. (2024). Smart customer service in unmanned retail store enhanced by large language model. *Scientific Reports*, 14, 71089. <https://doi.org/10.1038/s41598-024-71089-9>
- Wirtz, J., Kunz, W. H., Hartley, N., & Tarbit, J. (2023). How intelligent automation, service robots, and AI will reshape service products and their delivery. *Italian Journal of Marketing*, 2023, 289–312. <https://doi.org/10.1007/s43039-023-00076-1>
- Yuan, C., Liu, Y., Wang, S., & Zhang, T. (2025). Exploring technology, organization and environmental factors driving firms' generative AI adoption intention: The moderating role of trans-parasocial relation. *Journal of Business and Industrial Marketing*, 40(8), 1623–1640. <https://doi.org/10.1108/JBIM-04-2024-0257>
- Yue, B., Li, H., & Zhao, X. (2023). The impact of human-AI collaboration types on consumer evaluation and usage intention: A perspective of responsibility attribution. *Frontiers in Psychology*, 14, 1277861.

<https://doi.org/10.3389/fpsyg.2023.1277861>

Zhao, Z., Fan, W., Li, J., Liu, Y., Mei, X., Wang, Y., et al. (2024). Recommender systems in the era of large language models

(LLMs). *IEEE Transactions on Knowledge and Data Engineering*, 36(11), 6889–6907.
<https://doi.org/10.1109/TKDE.2024.3392335>