



Mapping the evolution of e-commerce research through co-word analysis: 2001–2020

Yang Bai^a, Hongxiu Li^{b,*}

^a School of Information Management, Central China Normal University, Wuhan, China

^b Department of Information and Knowledge Management, Faculty of Management and Business, Tampere University, Tampere, Finland

ARTICLE INFO

Keywords:

Co-word analysis
E-commerce, research trends
Research theme
Research evolution

ABSTRACT

This current study examines how research themes and trends have developed in the e-commerce field, aiming to provide a comprehensive understanding of e-commerce research based on prior literature. To answer this question, the current study applied a bibliometric method, word co-occurrence (co-word) analysis, to identify core research themes and trends in the e-commerce field, based on 17,416 author-provided keywords collected from 3,284 academic articles published between 2001 and 2020 in seven major journals in e-commerce field. Specifically, the data were analyzed based on two periods: 2001–2010 and 2011–2020. We found that the research themes in the e-commerce field have evolved alongside technology's development and diffusion. Some research themes, such as recommended system and eGovernment, have persisted as important research themes and studied continuously over the 20 years studied. Some research themes studied in 2001–2010, such as B2C and XML, disappeared in 2011–2020, whereas some new research topics, such as cloud computing, mobile technologies, artificial intelligence, and machine learning, emerged in recent years. This study provides a good understanding of e-commerce research's evolution, as well as useful suggestions to direct future research by scholars in the e-commerce field.

1. Introduction

In the past 20 years, we have witnessed e-commerce's rapid development alongside the technology development. e-Commerce has also attracted significant attention from researchers, and the research on e-commerce has increasingly grown, covering a broad range of topics. These topics' diversity has also raised questions about the core research themes in the e-commerce field, and scholars have called for studies examining the main research themes and evolution of e-commerce research. Such studies can provide knowledge about the core research themes in the e-commerce field and guide scholars' future research.

In response to this call, prior research has examined e-commerce research trends using different methods. For example, Wareham et al. (2005) applied a meta-analysis method to analyze some articles published in the e-commerce field, profiling e-commerce research by identifying the major research domains and the most salient themes in each of these domains. Meanwhile, Kauffman and Walden (2001) assessed the state-of-the-art knowledge about e-commerce phenomena, providing guidance for future work in this vein. Lin et al. (2016) reviewed papers published in the top 10 e-commerce journals from 1991 to 2014,

identifying the most influential articles, the connections between authors and these articles, and the main themes in e-commerce research. Moreover, Yoo and Jang (2019) applied content analysis to examine research themes for future e-commerce studies, based on over 1,000 articles published between 1987 and 2017 from three perspectives, including business models (BM), service relationships, and technology.

Some prior e-commerce studies have focused mainly on specific themes, using a fragmented approach that cannot fully represent a knowledge map of the e-commerce field. Additionally, though some prior studies have examined themes in e-commerce research, they have not addressed the research evolution or trends in e-commerce research vis-à-vis technology over time. Therefore, a study further examining the development of research themes and trends in this field generally, based on recent literature, is necessary to understand the evolution of e-commerce research. This examination should consider, for example, the main research themes in this field and how these themes have evolved alongside technology. Such a perspective can also provide a comprehensive knowledge map of e-commerce research for scholars.

To fill this gap in the literature, the current study examined research themes' evolution, based on 3,289 articles published in seven leading

* Corresponding author.

E-mail addresses: baiy@mails.ccnu.edu.cn (Y. Bai), Hongxiu.li@tuni.fi (H. Li).

<https://doi.org/10.1016/j.elerap.2022.101190>

Received 29 November 2021; Received in revised form 29 July 2022; Accepted 8 August 2022

Available online 17 August 2022

1567-4223/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

journals in the e-commerce field between 2001 and 2020, using word co-occurrence (co-word) analysis, aiming to answer the following three questions:

- i. What major themes emerged in e-commerce research from 2001 to 2020?
- ii. How have these major themes shaped e-commerce research?
- iii. How did these themes evolve through the studied 20-year period?

To answer these research questions, author-provided keywords applied in the published articles in seven major e-commerce journals in the two decades from 2001 to 2020 were collected and analyzed via co-word analysis method, and the major research topics and their evolution were identified and a visual knowledge map of e-commerce research over the past 20 years was presented.

The remainder of this paper is structured as follows. The second section reviews and discusses the literature, summarizing previous research in the e-commerce field and contextualizing this study's research context. Next, [Section 3](#) describes the research methods applied in this study. The fourth section presents and discusses our research findings. Finally, we address this study's contributions to the literature and discuss its limitations, as well as future research directions in e-commerce research.

2. Research background

In the late 1990s, the internet promoted the formation of e-commerce, and information technology's development in the intervening years has driven the rapid development of e-commerce over the past 20 years. For example, Web2.0 led to the emergence and development of the platform-based e-commerce model ([Yu, 2019](#)). Moreover, the emerging technologies—such as social media, cloud computing, big data, and artificial intelligence (AI)—have also scaled up e-commerce development ([Illumdeen, 2021](#)). e-Commerce has been driven by these technologies' rapid advancement, which might also have steered the evolution of research themes in this field.

Prior studies have applied various methods to understand research in the e-commerce field. For instance, [Ngai and Wat \(2002\)](#) presented a literature review and classification scheme for e-commerce research based on 275 journal articles published between 1993 and 1999 in nine e-commerce-related journals. They found exponential growth in research on the applications, technological issues, support and implementation, and other categories in the e-commerce field.

[Wareham et al. \(2005\)](#) studied 582 academic articles pertaining to e-commerce that had been published in academic journals between 1997 and 2003. With a meta-analysis method, they found that business-to-business (B2B), strategy, business-to-consumer (B2C), trust, and technology adoption were popular research topics in this field. Moreover, they found that survey was a dominant research method in the e-commerce field during the studied period. Their findings indicated great diversity in e-commerce research topics due to the field's interdisciplinary nature. Wareham et al. also pointed out that research topics in this field should have been developed alongside new technologies continuously applied to the e-commerce domain, providing a general understanding of e-commerce as an important research area in information systems (IS), management information systems (MIS) and international business research.

Based on 1,662 articles published in six leading journals of the e-commerce field between 2006 and 2010, [Shiau and Dwivedi \(2013\)](#) conducted multidimensional scaling, factor analysis, and cluster analysis of the research focus of these selected articles. They identified five core research themes in the e-commerce field, including e-markets, the acceptance and application of technology, the task-related application of e-commerce, identity and evaluation, and trust.

Additionally, [Lee et al. \(2007\)](#) conducted a diversity measurement analysis of articles across research disciplines. This analysis provided an

understanding of different journals' research focuses regarding e-commerce. Lee analyzed 1,103 articles that had been published across a total of 14 e-commerce, IS, and marketing journals between 1996 and 2005. This analysis found that IS and marketing journals tend to focus more on the mathematical modeling of e-commerce, whereas e-commerce journals tend to focus more on conceptual research in this area.

Some scholars have also applied bibliometric analysis methods to their research, providing a general understanding of research in the e-commerce field from different lenses, such as authors, articles, topics, and research trends. For example, based on 5,429 articles in the e-commerce field published from 1996 to 2015, [Tsai \(2015\)](#) found that the author productivity distribution across e-commerce research did not fit Lotka's law; rather, e-commerce closely relates to disciplines such as economics, engineering, computer science, and information science. Meanwhile, [Wang and Chen \(2010\)](#) conducted a statistical analysis and citation analysis of 4,948 articles published in the Science Citation Index Expanded (SCIE) and 2,875 articles published in the Social Science Citation Index (SSCI) from 1999 to 2008. They found that e-commerce research in SCIE focused more on computer and information science, whereas e-commerce articles in the SSCI database were more business-management-focused.

Moreover, [Lin et al. \(2016\)](#) reviewed articles about e-commerce published in the top 10 journals in the fields of IS and MIS from 1991 to 2014. They identified the most influential articles, the intellectual connection between authors and published articles, and the main research themes in e-commerce research via co-author analysis. [Du and Li \(2019\)](#) created a set of knowledge maps depicting, for example, co-citation, cited authors, institutions, and topics. Based on 1,130 articles with 25,502 associated references published in 2007–2018, they revealed the associated intellectual bases and hot research themes in mobile commerce (m-commerce) research. They found that m-commerce consumers, mobile technologies, m-banking (mobile banking) services, and mobile marketing were the major clusters in m-commerce research. Meanwhile, [Cui et al. \(2018\)](#) analyzed 12,089 papers in 503 publications between 2005 and 2017 via co-citation analysis. Based on CiteSpace software, they examined the knowledge structure of social commerce and its evolution by identifying key authors, articles, institutions, countries, and the major research topics.

Some research has also attempted to identify future e-commerce research trends by analyzing published articles in the field. For example, [Yoo and Jang \(2019\)](#) analyzed over 1,000 e-commerce articles published between 1987 and 2017. Using content analysis, they identified some promising research themes for future e-commerce research via a bibliographic classification, including BM, service relationships, and technology.

Additionally, a recent study by [Mou et al. \(2019\)](#) applied a bibliometric approach, using CiteSpace to assess keywords, authorship, reference articles, reference journals, institutions, and countries in e-commerce research. This analysis was based on 1,799 papers published in six e-commerce journals. The authors also identified research trends in the e-commerce field. In their study, [Mou et al. \(2019\)](#) identified social commerce, online reviews, social media, and word of mouth as future e-commerce research trends.

Previous studies have provided an understanding of research trends in the e-commerce research. However, some of these studies have focused on research themes in a specific field in e-commerce research, such as mobile commerce ([Du & Li, 2019](#)), and some studies have applied fragmented approach to examine the research themes based on identified major authors, institutions, and countries ([Lee et al., 2007](#); [Tsai, 2015](#)). Though some studies have studied the evolution of e-commerce research over time, these studies are either dated ([Ngai & Wat, 2002](#); [Wareham et al., 2005](#)), or the applied methods could not reveal the connections of research themes as well as the changes of research themes ([Mou et al., 2019](#); [Yoo & Jang, 2019](#)). There is a lack of comprehensive knowledge map of e-commerce research for scholars based on recent literature to reflect how the research themes and trends

Table 1
e-Commerce journals included in this study.

Journals	Starting year	Availability of keywords (year)	Articles	Keywords
International Journal of Electronic Commerce (IJEC)	1996	2000	408	2,360
Electronic Markets (EM)	1991	2000	554	3,382
Electronic Commerce Research and Applications (ECRA)	2002	2002	856	4,892
Electronic Commerce Research (ECR)	2001	2001	444	2,177
Journal of Electronic Commerce Research (JECR)	2000	2000	361	981
Journal of Theoretical and Applied Electronic Commerce Research (JTAECR)	2006	2006	303	1,697
Journal of Electronic Commerce in Organizations (JECO)	2003	2003	358	1,927
Total	—	—	3,284	17,416

develop with technology development over time. Therefore, it is necessary to investigate the evolution of e-commerce research based on recent literature.

3. Research methods

3.1. Co-word analysis

Co-word analysis is a bibliometric method that uses word co-occurrence analysis to identify subject areas that have characterized fundamental research at different periods (Callon et al., 1983), which “reveals patterns and trends in a specific discipline by measuring the association strengths of terms representative of relevant publications produced” in a research field (Ding 2001, p. 819). Co-word analysis can be conducted based on the co-occurrence network of keywords applied in published articles or of the words extracted from the title, the abstract, or the body of articles (Liu et al. 2016). Co-word analysis has also been applied to extract the important topics in a specific field based on large text data (Fronzetti Colladon, 2018).

Co-word analysis offers advantages compared to co-citation and co-author analysis by discovering connections between subjects in a specific research field while tracing a field’s scientific development and revealing knowledge in a specific field. It has been widely applied in prior research to investigate research topics in a specific field, based on published literature, by identifying relationships between research topics and these topics’ centrality to the field, as well as their internal structure (He, 1999), such as tourism (Hoz-Correa et al., 2018), e-learning (Bai et al., 2021), information retrieval (Ding et al., 2001), library and information science (Zong et al., 2013), and IS (Liu et al., 2016). In line with the prior literature, the current study applied co-word analysis to identify main research themes, knowledge structures, and research themes’ evolution in previous e-commerce studies.

In this study the co-word analysis is based on the keywords applied in published articles in e-commerce field. Specifically, each keyword in selected articles presents one node in the co-occurrence network. A co-occurrence is the frequency of two keywords both appear in the same article together (Ding 2001). The higher the co-occurrence frequency, the closer the relationship of these two keywords. Many co-occurrences of one keyword or pairs of keywords may indicate a research theme in e-commerce field (Ding 2001). A co-occurrence matrix is then built based on computing all the keywords as well as their connections in a structured network of e-commerce field, which helps identify the position of

various research themes and the research trends that have evolved over time in e-commerce field.

3.2. Data collection

In this study, we collected the author-provided keywords included in the articles published in the seven major e-commerce journals in the “Business, Management, and Accounting” category of the SCImago Journal Rank (SJR) 2019. The details of the seven journals are presented in Table 1. Most of the author-provided keywords in the articles published between 2001 and 2020 were extracted from the article webpages in the online database of the seven journals via self-developed tools. For some articles published at a relatively earlier time without keywords available from the article webpages we manually collected the keywords from the full-text article documents in PDF. In total, 17,416 keywords were collected from 3,284 articles. Table 1 presents more details of the distribution of keywords in the seven journals. The keywords with frequencies lower than 3 were excluded in this study since few researchers have applied them in their articles to represent the topics of these articles. Thus, ultimately, 5,792 of the identified 17,416 keywords were selected in this study.

To obtain more accurate outcomes and guarantee a diversity of topics, we standardized the collected keywords manually. For example, we merged synonyms (e.g.: *online market* and *electronic market*; *customer satisfaction* and *user satisfaction*), the singular and the plural forms of different gerunds and nouns (e.g., *website* and *websites*), and abbreviations (e.g., *word of mouth* and *WOM*; *user-generated content* and *UGC*). We also filtered out some keywords that were too general and did not represent any research topics (e.g., *general*, *online*). Additionally, we excluded the *e-commerce* keyword and its variations (e.g., *electronic commerce*) since *e-commerce* is a common keyword applied to e-commerce articles. Two authors conducted this keyword merging task.

3.3. Data analysis methods

This study applied different data analysis methods to support its co-word analysis. We first identified the topic structure of the selected articles by constructing a keyword matrix. Then, we examined the selected articles’ research themes, based on clustering analysis. Next, we explored the knowledge networks and the structural features of the identified research themes via Social Network Analysis (SNA) to understand the status and future development potential of these research themes, based on a strategic diagram. The following subsections present these methods in more detail.

3.3.1. Constructing keyword matrix

A research field’s topic structure can be obtained from the content’s relationships to the literature on the topic, such as topics’ affiliation or distribution reflected in keyword-paper relationships, or strong or weak research themes presented by keyword-to-keyword co-occurrence relationships. To explore the thematic structure of e-commerce research over the past 20 years, we set up a keyword-paper matrix and a co-occurrence matrix via Python programming, based on latent semantic analysis (LSA).

According to LSA, keywords can denote the main semantics of an article, whereas the relationships between the keywords in a collection of articles can represent a set of articles’ knowledge structures (Laudauer & Dumais, 1997). In other words, such knowledge structures can be detected by mapping keywords in an article vector space, based on keywords’ distribution in a set of articles. Thus, classes of keywords can be obtained by clustering keyword vectors based on article positions in a given field. The keyword co-occurrence matrix is obtained from the keyword-paper matrix, which is constructed by statistically counting whether the keywords appear in each article, and based on that, each two keywords are further counted to form the keyword co-occurrence matrix by how many articles they appear in together. Based on a co-

occurrence matrix, patterns of keyword co-occurrence help identify highlighted research topics in a field (Liu & Mei, 2016), and the knowledge networks and structural features of a specific research field can be further examined via SNA and strategy maps.

3.3.2. Clustering analysis

In LSA, a research topic can be represented by a set of closely related keywords, based on hierarchical cluster analysis. Hierarchical cluster analysis constructs classes of keywords by continuously merging the closest pairs of keywords, measuring the degree of relationships between keywords by calculating the similarity (or dissimilarity) between target keywords, and then aggregating the keyword pairs via a clustering algorithm.

In the current study, the Bray-Curtis dissimilarity algorithm was employed to evaluate the similarity (or dissimilarity) between keywords based on their connection relationships and to get a list of all the keywords in a research cluster arranged by similarity. This algorithm is robust to nonlinearity, and it can be applied to analyze binary or categorical data, as well as data with multiple zero-valued dimensions (Beals, 1984; Field et al., 1982). It has been employed in different clustering analyses in the prior literature (Bai et al., 2021; Jayaprakash et al., 2016). We also applied Ward's method (Ward, 1963) to further hierarchically merge similar keywords until these keywords are clustered into a single cluster, using Python's SciPy package for coalescent clustering. The clustering analysis helps to detect the different clusters and the popular research topics in each cluster in e-commerce research in the period we studied.

3.3.3. Social network analysis (SNA)

The SNA method has been widely applied in bibliometric research to explore knowledge structures in specific research fields since it can examine topics' internal structure and characteristics from a network view. Core-periphery, structural hole, degree centrality and network density in SNA were applied in this study to explore the knowledge structure of e-commerce topics via a network package NetworkX in Python.

Core-periphery analysis in SNA can help identify the core nodes densely connected with each other as well as the peripheral nodes which are connected to the core nodes but sparsely connected to each other in a network (Borgatti & Everett, 1999). In other words, the core nodes are well connected with the other core nodes and the peripheral nodes, but the periphery nodes are loosely connected with each other. The core and periphery values in this study are calculated based on the ties each keyword (topic) set up with other keywords (topics) in the co-occurrence network. The values in core-periphery analysis can help detect the fundamental research themes in e-commerce field.

Structural hole analysis can calculate the degree of a non-redundant connection between two disconnected nodes (Burt, 1992). For example, in a co-occurrence network, keyword C is connected to keywords A and B, but keyword A is not connected to keyword B. Therefore, keyword C will have a high structural hole, acting as a bridge in the network to build the bridge between the two unconnected keywords A and B. If a keyword is more connected with other isolated keywords in a network, this keyword will have a higher structural hole value. The value of the structural hole is calculated based on the quantity of alters minus the average degree of alters within the co-occurrence network via the NetworkX package in Python (Burt, 1992). The structure hole analysis can help identify the bridging research topics in e-commerce field.

Moreover, we calculated degree centrality and network density through SNA to examine the degree of association between different research themes formed by some keywords (nodes in a network). Degree centrality estimates the degree to which a research theme is connected with other themes in a network (Nielsen & Thomsen, 2011). The degree centrality of a cluster can be described by the number of edges connected to other clusters in the co-occurrence network. The higher the degree, the more central the cluster is. Specifically, the total number of

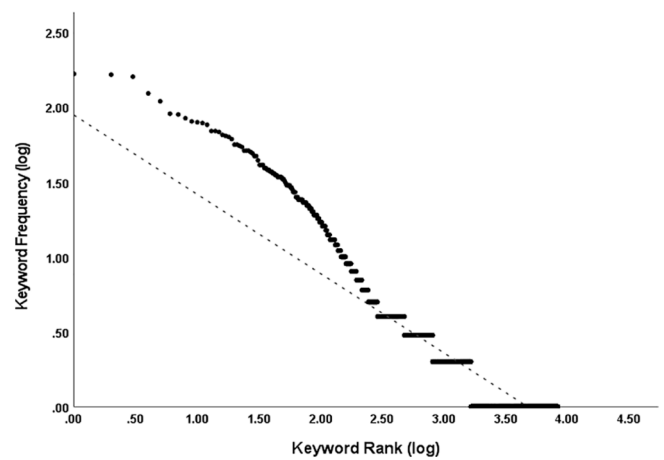


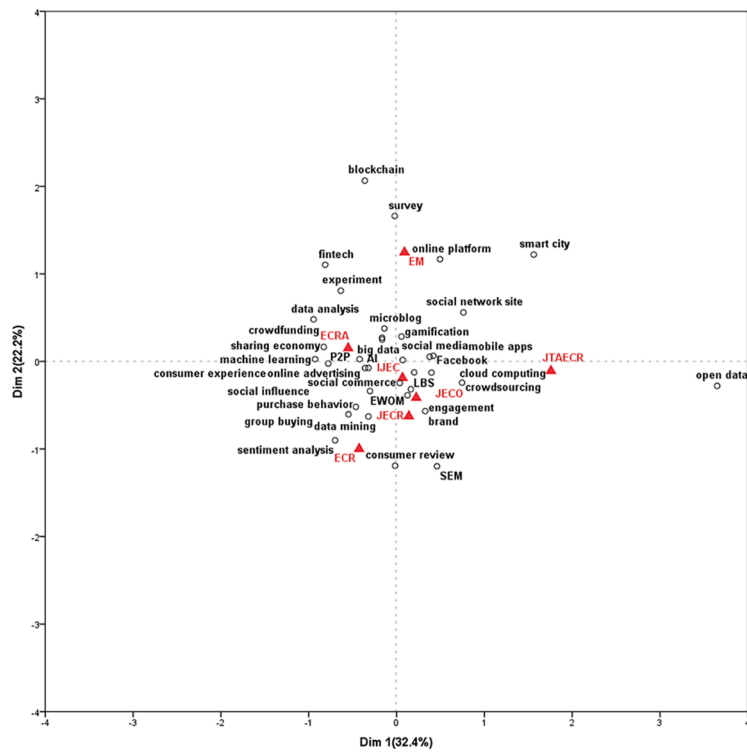
Fig. 1. The power-law distribution of keyword frequencies.

ties is divided by the maximum number of possible ties to calculate the density value. If the connections between a research cluster and the other clusters are strong, it indicates that this research cluster is essential to the research network (Callon et al., 1991). Meanwhile, network density measures a node's strength in a network, describing the degree to which a subnet or network is internally connected (He, 1999). The density of a cluster can be calculated by the ratio of the total number of relationships of all nodes constituting the cluster to the theoretical maximum number of relationships of these nodes, which can be used to describe the intensity of inter-connections within the cluster in the co-occurrence network. Density's value can explain a research theme's ability to remain consistent or evolve over time in a network. In this study, the centrality and density were computed via the software UCINET.

Combining centrality and density values in a network allows for the creation of two-dimensional strategic diagrams with x-axes reflecting centrality and y-axes reflecting density (Bauin et al., 1991). Specifically, in this study the centrality of a research theme reflects the extent to which the research theme is connected with other research themes, and its value reflects the intensity of the connections. Moreover, the density of a research theme reflects the extent to which the research theme is internally connected in a network. In a strategic diagram, the research themes located in Quadrant I can be regarded as motor themes in a research field, which have high density and centrality and are central to the research field. Meanwhile, the themes in Quadrant II are interconnected but not closely linked externally to other themes, which are more specialized and peripheral to the motor themes and might constitute emerging research themes currently under development. The themes located in Quadrant III are relatively marginalized and peripheral, representing fading or emerging themes in a research field. Finally, the research themes in Quadrant IV are weakly structured and underdeveloped, with the potential to evolve as significant research themes in a field (Muñoz-Leiva et al., 2012). Thus, a strategy diagram can describe each research domain's position in an entire research field, as well as these themes' statuses and potential. The network centrality and density values of each topic in co-occurrence network calculated via SNA help create the two-dimensional strategic diagrams of each research topic in e-commerce research in this study and identify the cohesiveness and maturity of each research cluster in the e-commerce knowledge network over the past 20 years.

4. Results

The keyword frequency distribution in a network can help to get insight into the state of the research topics of a research field. Rank-frequency relationships are often combined with Zipf's law (Zipf,



Note: Δ in red: journals; \circ : research topics.

Fig. 3. Keywords distribution in the seven studied e-commerce journals.

emerging topics in e-commerce research during 2011–2020. Therefore, these topics gained importance in e-commerce research (Javid et al., 2019; Norian et al., 2020).

Further, we conducted a correspondence analysis of the 36 new keywords that appeared between 2010 and 2020 in the selected journals, exploring the different e-commerce journals’ concerns with new research topics. As Fig. 3 shows, most of these 36 new keywords were located at the center and near the seven e-commerce journals. This placement indicated that the seven studied e-commerce journals contributed to the development of new research topics that emerged in 2010–2020.

4.1. Popular, core, and bridging research topics

In a co-word network, popular research topics mean that the research topics (nodes) have high frequency in comparison with other research topics in the co-occurrence network; core research topics are the research topics that constitute the densest connections in the co-occurrence network, which were identified based on the core-periphery analysis; while the research topics (nodes) with high structural hole values represent a bridging role due to the indirect connection with other research topics (nodes). By identifying the popular, core, and bridging research themes in the two different time periods, we can understand the position and the importance of different research topics in e-commerce field in a network structure.

We identified the popular, core, and bridging research topics in e-

Table 2
Popular, core, and bridging research topics in 2001–2010.

No.	Popular research topics	Frequency	Core research topics	Core value	Bridging research topics	Structure hole value
1	marketing	92	service quality	0.352	marketing	45.854
2	auction	72	trust	0.339	trust	41.284
3	trust	69	marketing	0.292	service	40.925
4	service	56	auction	0.228	auction	28.767
5	mobile e-commerce	46	service	0.22	SCM	28.412
6	internet	44	satisfaction	0.21	consumer behavior	27.684
7	consumer behavior	41	consumer behavior	0.188	strategy	26.373
8	SCM	35	information system	0.182	internet	25.155
9	online banking	34	online shopping	0.17	economic analysis	23.222
10	privacy	31	strategy	0.17	B2C	20.535
11	eGovernment	30	privacy	0.16	technology adoption	19.921
12	online shopping	30	retailing	0.153	retailing	19.812
13	case study	29	pricing	0.142	pricing	18.87
14	retailing	29	SCM	0.139	online banking	18.731
15	strategy	29	mobile e-commerce	0.139	value	18.25
16	pricing	28	B2C	0.133	B2B	18.247
17	B2B	27	reputation	0.127	software agent	18.176

Note: The text in boldface refers to the keywords that appeared three times within the popular, core, and bridging research topics.

Table 3
Popular, core, and bridging research topics in 2011–2020.

No.	Popular research topics	Frequency	Core research topics	Core value	Bridging research topics	Structure hole value
1	online review	143	online review	0.385	trust	52.84
2	trust	95	consumer review	0.327	online review	48.092
3	social media	80	social media	0.278	consumer behavior	47.003
4	marketing	74	trust	0.26	marketing	45.074
5	consumer behavior	68	service quality	0.246	social media	43.71
6	social commerce	61	social commerce	0.225	<i>value</i>	38.976
7	online banking	55	consumer behavior	0.197	big data	36.138
8	<i>value</i>	52	online shopping	0.183	business model	34.791
9	action	51	<i>social network</i>	0.159	online shopping	34.272
10	retailing	50	reputation	0.158	social commerce	33.862
11	<i>privacy</i>	47	data mining	0.151	<i>mobile e-commerce</i>	33.411
12	<i>recommendation system</i>	47	social network site	0.148	case study	32.503
13	online shopping	46	<i>recommendation system</i>	0.14	<i>privacy</i>	31.338
14	<i>satisfaction</i>	45	<i>satisfaction</i>	0.139	empirical research	30.96
15	<i>mobile e-commerce</i>	44	marketing	0.135	<i>social network</i>	30.632

Note: The text in boldface refers to the keywords that appeared three times within the popular, core, and bridging research topics.

commerce research during 2001–2010 and 2011–2020, respectively. While we identified 17 keywords as core research topics in 2001–2010 based on the method of core-periphery analysis, with a concentration value of 0.826, and 15 keywords with a concentration value of 0.848 in 2011–2020. The concentration value refers to the degree to which a set of keyword nodes representing the degree of core components in a co-occurrence network can represent the overall core components of the co-occurrence network, which can be calculated based on the core–edge analysis. We also identified popular research topics based on the co-occurrence frequency of keywords that appeared in collected articles, and bridging topics based on keywords’ structural hole values in each period. More details are presented in tables 2 and 3.

As Table 2 shows, in 2001–2010, most keywords appeared across these three categories (*popular*, *core*, and *bridging* research topics), suggesting a high degree of consistency in research themes throughout 2001–2010. Specifically, nine research themes—specifically, *marketing*, *action*, *trust*, *service*, *consumer behavior*, *SCM*, *retailing*, *strategy*, and *pricing*—were popular, core, and bridge topics, implying these topics’ importance to the e-commerce field from 2001 to 2010. Additionally, economic analysis, technology adoption, value and software agent were found to have low popularity despite being bridge topics, indicating that although these four topics were not core-relevant to the knowledge network of e-commerce research during 2001–2010, they were relevant to some isolated research during this period. We also found that some research topics—such as *online banking*, *internet* and *B2B*—were not core topics during 2001–2010 but were popular and bridging topics during that period. This finding shows that although these themes were popular, the links between them were weak, and they have not yet become core themes in the e-commerce field.

In the 2011–2020 period, 10 of the 15 keywords were identified as both popular and bridging topics—specifically, *online review*, *trust*, *social media*, *marketing*, *consumer behavior*, *social commerce*, *value*, *privacy*, *online shopping*, and *mobile e-commerce*. This finding indicates that these bridge topics have attracted scholarly attention and maintained cohesiveness across e-commerce research. Notably, seven keywords—particularly, *online review*, *trust*, *social media*, *marketing*,

consumer behavior, *social commerce*, and *online shopping*—were identified as the popular, core, and bridging themes during 2011–2020. This finding shows these topics’ importance and the high consistency of e-commerce research during that period.

As Tables 2 and 3 show, though the topics of *economic analysis*, *technology adoption*, *software agents*, *big data*, *BM*, and *empirical research* were not identified as popular or core topics, they had high structural hole values in 2001–2020. These topics are closely related to research methodology and technology in e-commerce, and they play an important role in bringing those relatively isolated themes together to address some issues in this field. Some topics—such as *the internet*, *online banking*, *B2B*, *value*, *privacy*, and *mobile e-commerce*—were identified as popular and bridging topics but not as core topics in the 20 years studied. This finding suggests that the research themes are popular but relatively general topics compared to topics identified as core themes; nevertheless, they help condense other discrete themes in e-commerce research during 2001–2020. *Service quality*, *information system*, *reputation*, *consumer review*, *data mining*, and *social network site* were found to be core topics, but they were not popular or bridging topics, indicating that these topics are relatively traditional in the e-commerce field.

We also found that three keywords—specifically, *online shopping*, *mobile e-commerce*, and *privacy*—had evolved as new bridging topics during 2011–2020 from popular and core topics in 2001–2010. This development suggests that researchers had pushed these three topics from mainstream research themes to strong cohesive research themes in 2001–2020. Meanwhile, some core, and bridging topics in 2001–2010 had lost their core, and bridging status in 2011–2020, such as *action*, *service*, *SCM*, *retailing*, *strategy*, *pricing*, and *B2C*. These findings indicate that these topics are losing their importance to e-commerce research over the past 20 years.

Furthermore, we found that bridging themes changed considerably in e-commerce research during 2001–2020. For example, 13 keywords in this period disappeared in 2011–2020, and 11 new keywords with high structural hole values (bridging topics) appeared in 2011–2020. Such changes also occurred among the core and popular topics we identified, such as 10 keywords in 2001–2010 that disappeared in

Table 4
Disappearing and emerging popular, core and bridging research topics based on the comparison of research in 2001–2010 and 2011–2020.

keywords state	Popular research topics	Core research topics	Bridging research topics
Disappear in 2011–2020	SCM, service, eGovernment, case study, internet, strategy, pricing, and B2B	Auction, service, information system, strategy, privacy, retailing, pricing, SCM, mobile e-commerce, and B2C	Service, auction, SCM, strategy, internet, economic analysis, B2C, technology adoption, retailing, pricing, online banking, B2B, and software agent
New in 2011–2020	Online review, social commerce, social media, value, recommendation system, and satisfaction	Online review, consumer review, social media, social commerce, social network, data mining, social network site, and recommendation system	Online review, social media, big data, business model, online shopping, social commerce, mobile e-commerce, case study, privacy, empirical research, and social network

Table 5
Identified research clusters in e-commerce research for 2001–2010.

Cluster (size)	Keywords
A1 (8)	eContract; eGovernment; interoperability; matchmaking; ontology; semantic web; service; XML
A2 (3)	business network; healthcare; procurement
A3 (9)	agent; commitment; evaluation; multiagent system; negotiation; RFID; SCM; simulation; software agent
A4 (4)	electronic cash; micro payment; network effect; software
A5 (7)	business process; economic analysis; online banking; risk; standardization; TAM; technology adoption
A6 (8)	adoption; cross-border; developing country; innovation diffusion; internet; performance; small business; SME
A7 (5)	economics; fairness; game theory; intelligent agent; mechanism design
A8 (7)	accessibility; auction; cryptography; eBay; reputation; security; trust
A9 (2)	clustering; recommendation system
A10 (6)	B2C; loyalty; online shopping; retailing; satisfaction; service quality
A11 (9)	B2B; cooperation; culture; EDI; empirical research; information system; knowledge management; usability; website
A12 (6)	algorithm; decision support; elaboration likelihood model; online review; personalization; privacy
A13 (24)	business model; case study; China; consumer behavior; CRM; digital rights management; ICT; information goods; innovation; interactivity; intermediation; marketing; mobile e-commerce; mobile service; mobile technology; online community; pricing; search; social network; strategy; ubiquitous computing; value; virtual world; web2.0

Note: The keywords in boldface refer to the keywords with a high frequency, which represent these clusters.

Table 6
Identified research clusters in e-commerce research for 2011–2020.

Cluster (size)	Keywords
B1 (12)	brand; EWOM; Facebook; microblog; online community; social commerce; social influence; social media; social network; social network site; Twitter; web2.0
B2 (3)	auction; mechanism design; procurement
B3 (8)	cloud computing; game theory; online advertising; pricing; retailing; SCM; search; strategy
B4 (2)	privacy; security
B5 (5)	eGovernment; ICT; information system; internet; interoperability
B6 (7)	B2C; marketing; mobile e-commerce; network effect; online payment; risk; service
B7 (14)	consumer behavior; consumer experience; loyalty; online shopping; personalization; purchase behavior; reputation; satisfaction; SEM; service quality; sharing economy; survey; TAM; trust
B8 (4)	blockchain; fintech; innovation; online platform
B9 (8)	engagement; enjoyment; gamification; innovation diffusion; mobile apps; mobile technology; technology adoption; virtual world
B10 (21)	adoption; AI; business model; case study; CRM; crowdfunding; crowdsourcing; decision support; experiment; group buying; internet of things; LBS; mobile service; online banking; open data; P2P; product review; smart city; social capital; software; value
B11 (7)	algorithm; big data; China; data analysis; empirical research; machine learning; website
B12 (2)	clustering; recommendation system
B13 (5)	consumer review; cross-border; data mining; online review; sentiment analysis

Note: The keywords in boldface refer to the keywords with high frequency, which represent these clusters.

2011–2020 and eight new keywords that appeared in 2011–2020 as core topics, eight keywords in 2001–2010 disappeared in 2011–2020, and six new keywords that were added in 2011–2020 as popular topics. These findings suggest that research themes in e-commerce research have changed considerably over the last 20 years. The addition of *online review*, *social media*, and *social commerce* as popular, core, and bridging topics has driven e-commerce research forward. Meanwhile, *marketing*, *trust*, and *consumer behavior* were maintained as popular, core, and

bridging research topics in both 2001–2010 and 2011–2020, suggesting that these topics are classic and significant to e-commerce research. More details on the change of popular, core, and bridging research topics in the two researched periods are presented in [Table 4](#).

4.2. Major research topics and their status in e-commerce research

We constructed a keyword-paper co-occurrence correlation matrix via Python’s SciPy package. The distances between different research topics were calculated based on the Bray-Curtis heterogeneity algorithm (Bray & Curtis, 1957), and the topics were also classified into different research clusters via Ward’s method (Ward, 1963). More details on the cluster analysis can be found in Appendix [Fig. A1](#) and [Fig. A2](#). Finally, 13 clusters were returned for the 2001–2010 period (labeled A1–A13; see [Table 5](#)), as well as 13 clusters for the 2011–2021 period (labeled B1–B13; see [Table 6](#)). Each cluster represents a research subfield in the wider field of e-commerce. For example, A9 relates to the recommender systems while A10 pertains to B2C online shopping (see [Table 5](#)), and B1 relates to social commerce while B4 pertains to privacy and security in e-commerce (see [Table 6](#)).

The structural indicators of each research cluster helped explain the status and importance of these research subfields to e-commerce research. We calculated the values for total keyword frequency, average keyword frequency, total co-word frequency, and average co-word frequency for all research clusters to measure the popularity of each research cluster. As [Tables 7 and 8](#) show, clusters A8 (*accessibility, auction, cryptography, eBay, reputation, security, trust*) and A10 (*B2C, loyalty, online shopping, retailing, satisfaction, service quality*) from 200 to 2010 and clusters B12 (*clustering, recommendation system*) and B13 (*consumer review, cross-border, data mining, online review, sentiment analysis*) from 2011 to 2020 were detected as the most popular research subfield in e-commerce research in their respective periods since these clusters had the highest values for average keyword frequency as well as average co-word frequency. The findings suggest that these topics included in the four research clusters were popular research themes that had attracted significant attention from scholars in 2001–2010 and 2011–2020, respectively.

4.3. Research themes’ maturity and cohesion

To evaluate the cohesiveness and maturity of each research cluster in the e-commerce knowledge network over the past 20 years, we used the SNA method to calculate values for degree centrality and network density (See [tables 7 and 8](#)) for all research clusters in the co-occurrence network over the two research decades. Based on the centrality and density values in the co-occurrence network, a strategy map is constructed to visualize the strategic location of each research cluster in the co-occurrence network for periods 2001–2010 and 2011–2020 respectively (See [Fig. 4a and 4b](#)). The x-axis represents the centrality of a research topic in the co-occurrence network, reflecting external cohesion, and the y-axis represents the density of a research topic in the co-occurrence network, measuring the internal maturity.

A strategic diagram of the research themes in 2001–2010 is presented in [Fig. 4a](#). Clusters A1 (*eContract, eGovernment, interoperability, matchmaking, ontology, semantic web, service, XML*) and A5 (*business process, economic analysis, online banking, risk, standardization, TAM, technology adoption*) hold in Quadrant I. They have relatively high centrality and density, suggesting that these research clusters have a good level of internal maturity and are widely connected with other research clusters. Thus, Clusters A1 and A 5 are core research themes in the e-commerce research for the period 2001–2010.

Meanwhile, clusters A2 (*business network, healthcare, procurement*), A4 (*electronic cash, micro payment, network effect, software*), A9 (*clustering, recommendation system*), and A10 (*B2C, loyalty, online shopping, retailing, satisfaction, service quality*) are located in Quadrant II. The density values for these research clusters are high and their centrality

Table 7

Structural indicators of the 2001–2010 research clusters.

Cluster	Total frequency for keyword	Total frequency for co-word	Average frequency for keyword	Average frequency for co-word	Centrality	Density
A1	154	231	19.250	28.875	0.967	0.750
A2	29	49	9.667	16.333	0.463	1.000
A3	174	255	19.333	28.333	0.787	0.611
A4	29	35	7.250	8.750	0.553	0.667
A5	117	169	16.714	24.143	0.736	0.714
A6	152	167	19.000	20.875	0.700	0.607
A7	40	55	8.000	11.000	0.505	0.500
A8	203	270	29.000	38.571	0.648	0.619
A9	33	35	16.500	17.500	0.260	1.000
A10	126	202	21.000	33.667	0.511	0.800
A11	126	169	14.000	18.778	0.596	0.417
A12	93	102	15.500	17.000	0.402	0.400
A13	494	549	20.583	22.875	0.838	0.279
Average	136.154	176.000	16.600	22.054	0.613	0.643

Table 8

Structural indicators of the 2011–2020 research clusters.

Cluster	Total frequency for keyword	Total frequency for co-word	Average frequency for keyword	Average frequency for co-word	Centrality	Density
B1	367	653	30.583	54.417	1.000	0.697
B2	75	70	25.000	23.333	0.642	1.000
B3	214	205	26.750	25.625	0.900	0.500
B4	68	100	34.000	50.000	0.375	1.000
B5	95	118	19.000	23.600	0.753	0.800
B6	208	273	29.714	39.000	0.835	0.524
B7	468	735	33.429	52.500	0.940	0.648
B8	67	113	16.750	28.250	0.511	1.000
B9	179	242	22.375	30.250	0.811	0.750
B10	454	535	21.619	25.476	0.974	0.262
B11	155	279	22.143	39.857	0.593	0.857
B12	87	113	43.500	56.500	0.260	1.000
B13	242	324	48.400	64.800	0.355	1.000
Average	206.077	289.231	28.713	39.508	0.688	0.772

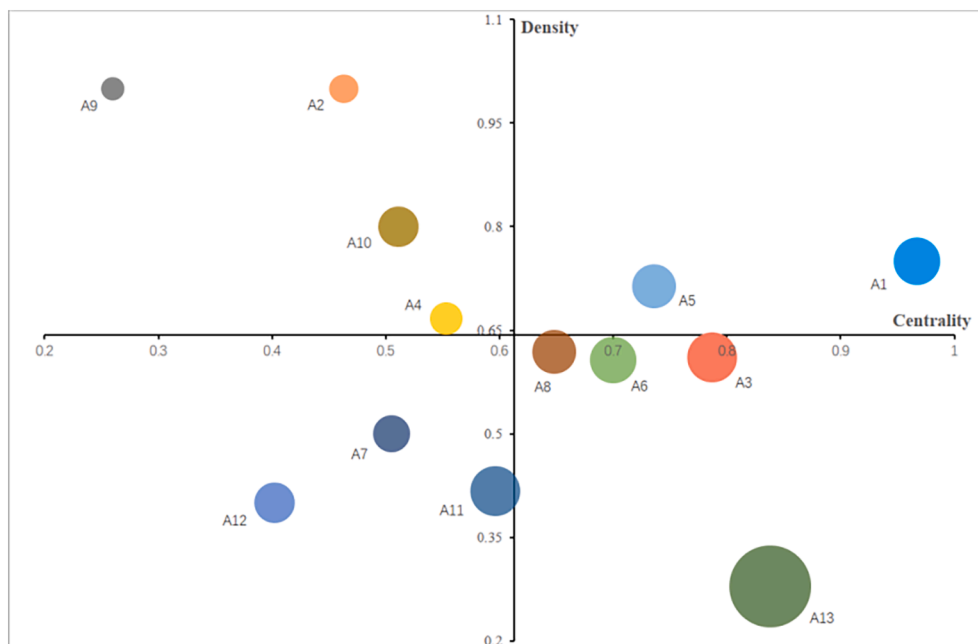


Fig. 4a. Strategic diagrams for e-commerce research during 2001–2010.

values are low, indicating that the research topics in the research clusters A2, A4, A9, and A10 are assumed to have been well developed and be mature but these topics are isolated from other research topics in e-commerce field in 2001–2010.

Clusters A7 (*economics, fairness, game theory, intelligent agent,*

mechanism design), A11 (*B2B, cooperation, culture, EDI, empirical research, information system, knowledge management, usability, website*), and A12 (*algorithm, decision support, elaboration likelihood model, online review, personalization, privacy*) are located in Quadrant III. Their centrality and density are low. This finding means that the research topics in these

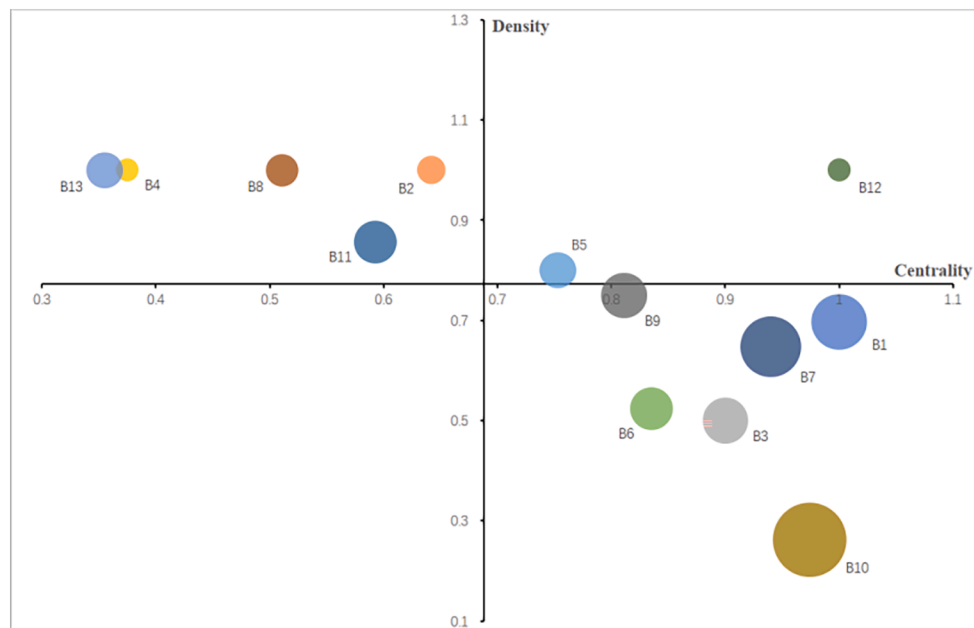


Fig. 4b. Strategic diagrams for e-commerce research during 2011–2020.

three clusters (A7, A11, and A12) have loose connections with each other, and also share minimal links with other clusters in the knowledge network. Thus, the research topics in A7, A11, and A12 are assumed to change. Such as these topics might fade away or develop as new core research topics in the e-commerce field.

Additionally, clusters A3 (*agent, commitment, evaluation, multiagent system, negotiation, RFID, SCM, simulation, software agent*), A6 (*adoption, cross-border, developing country, innovation diffusion, internet, performance, small business, SME*), A8 (*accessibility, auction, cryptography, eBay, reputation, security, trust*), and A13 (*business model, case study, China, consumer behavior, CRM, digital rights management, ICT, information goods, innovation, interactivity, intermediation, marketing, mobile e-commerce, mobile service, mobile technology, online community, pricing, search, social network, strategy, ubiquitous computing, value, virtual world, web 2.0*) are located in Quadrant IV. Their values for centrality are high but the values for density are low, suggesting that A3, A6, A8, and A13 could become core research themes despite their underdevelopment in 2001–2010. Furthermore, clusters A3, A6, A8, and A13 have high total keyword frequency values, implying that the research topics in these four clusters have attracted significant scholarly attention and been developed as more cohesive research subfields of the e-commerce field.

In the 2011–2020 period, the 13 research clusters are spread over quadrants I, II, and IV, as Fig. 4b shows. Specifically, clusters B5 (*eGovernment, ICT, information system, internet, interoperability*), and B12 (*clustering, recommendation system*) are located in Quadrant I as core research themes in e-commerce research during 2011–2020. Their high centrality and high density suggest that the research topics in these two clusters are well developed and widely connected to each other.

Clusters B2 (*auction, mechanism design, procurement*), B4 (*privacy, security*), B8 (*blockchain, fintech, innovation, online platform*), B11 (*algorithm, big data, China, data analysis, empirical research, machine learning, website*), and B13 (*consumer review, cross-border, data mining, online review, sentiment analysis*) are located in Quadrant II. They are mature yet marginal research themes since their topics are closely linked internally (high density) but isolated from each other (low centrality).

Meanwhile, clusters B1 (*brand, EWOM, Facebook, microblog, online community, social commerce, social influence, social media, social network, social network site, Twitter, web 2.0*), B3 (*cloud computing, game theory, online advertising, pricing, retailing, SCM, search, strategy*), B6 (*B2C, marketing, mobile e-commerce, network effect, online payment, risk, service*), B7

(*consumer behavior, consumer experience, loyalty, online shopping, personalization, purchase behavior, reputation, satisfaction, SEM, service quality, sharing economy, survey, TAM, trust*), B9 (*engagement, enjoyment, gamification, innovation diffusion, mobile apps, mobile technology, technology adoption, virtual world*), and B10 (*adoption, AI, business model, case study, CRM, crowdfunding, crowdsourcing, decision support, experiment, group buying, internet of things, LBS, mobile service, online banking, open data, P2P, product review, smart city, social capital, software, value*) are located in Quadrant IV. Our results suggest that the research topics included in these six clusters are core—yet undeveloped—topics in the e-commerce field from 2011 to 2020.

Additionally, we found that some clusters—such as A13, B1, B7, and B10—included much more keywords in the e-commerce field compared to other clusters. Though clusters A1 and B3 have fewer keywords, they are the core research themes in their respective periods, and they have attracted research interest in some traditional research topics, such as *eContract, eGovernment, and retailing*.

The strategic diagrams of the research themes in the two time periods help get an understanding of the center of e-commerce research in the 20 years we studied. Such as clusters A1, A3 and A13 (see Fig. 4a) as well as B1, B10 and B12 (See Fig. 4b) are found to be the center of e-commerce research in 2001–2020 since these research clusters have the highest centrality values. The research topics included in these research clusters cover various technologies and theories in e-commerce research, indicating that technology has been the core in e-commerce research to understand the impact of e-commerce on individuals, business, and society from different theoretical views.

Furthermore, we noticed the change of the center of e-commerce research in the 20 years since some keywords related to technology disappeared and some new technology keywords emerged over time. These findings indicate that the development of e-commerce has relied on technological development at different times, and scholars' attention has also shifted to some innovative technologies but not to mature technologies. For example, *XML* in Cluster A1 lost its importance to e-commerce research in 2011–2020 due to rapid technological development. Meanwhile, new technologies—such as *social media, microblog, online community, social network site, and web 2.0* from B1, and *crowdsourcing, internet of things; and LBS* from B10—have attracted researchers' attention. As Fedorko et al. (2018) stated, the technologies in the e-commerce field keep changing, and importantly, e-commerce

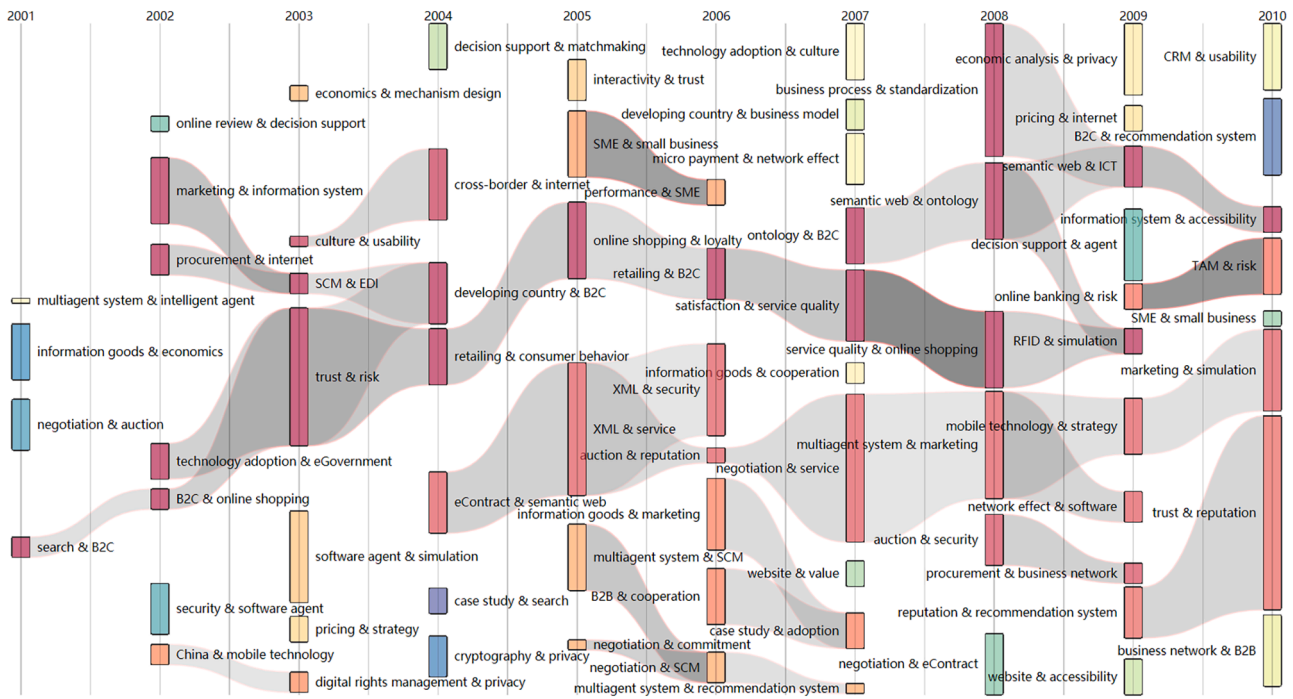


Fig. 5a. The evolution of e-commerce research in 2001–2010.

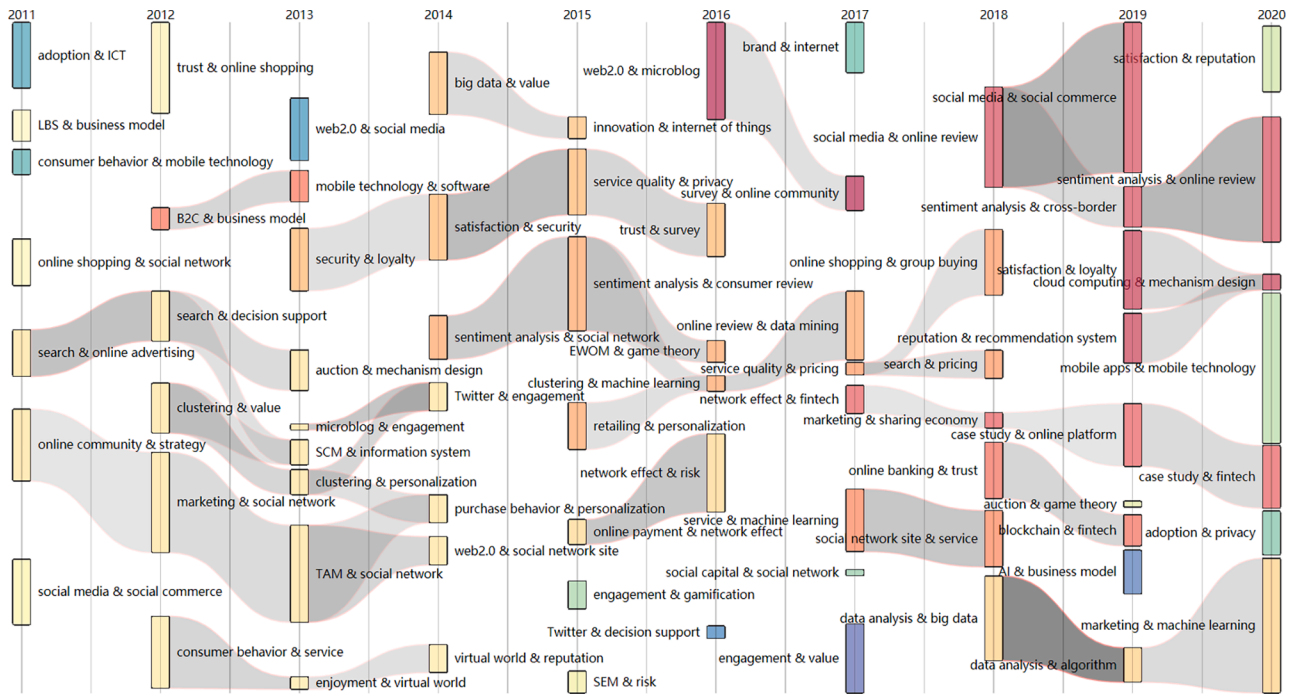


Fig. 5b. The evolution of e-commerce research in 2011–2020.

research must keep up with the field’s rapid technological advances.

4.4. The evolution of e-commerce research themes

To understand how the themes in e-commerce research evolved in the 20 years that we studied, we employed a visualization tool CorText (<https://www.cortext.net>) to visualize the research topics’ details and trajectories during the two examined periods (See Fig. 5a and 5b). In Fig. 5a and 5b, each block represents a core research topic in a year, which describes the most connected relationships between nodes in the

co-occurrence network. The higher the length of a block, the more popular the topic formed based on the connected relationships of nodes in a year. The tube connecting different blocks across two different years describes the number of co-occurrences between the two blocks and other nodes. The wider a tube is, the more the research topics in a block in the previous year shares connections with other themes when it evolves to another research topic in the following year. We generated evolution maps for these topics during these two studied periods (Leydesdorff & Goldstone, 2014).

We found two main subdivisions in e-commerce research during

2001–2010 (See Fig. 5a). The subject in the first subdivision mainly focused on the development of B2C e-commerce and related issues in e-commerce development. B2C maintained strong continuity during this period. Though research on the topic was weak in the e-commerce field in 2001–2002, it was well developed from 2002 to 2009 because it absorbed other research topics (e.g., *trust*, *consumer behavior*, *loyalty*, *service quality*). The subject's development also helped expand a new subdivision in e-commerce research through integration with the *ontologies*, *semantic web*, and *information systems* topics from 2007 to 2010. Another subdivision reflects marketing and strategies in e-commerce. In 2004, *eContract* attracted significant research attention, and this subject was developed by further integrating some other related topics (such as *negotiation*, *multiagent system*, *marketing*, and *strategy*) in 2005–2010. During this period, various topics related to e-commerce technologies emerged, such as *multiagent system* and *intelligent agent* in 2001, *internet* in 2002, *SCM* and *EDI* in 2003, and *RFID* in 2008. However, these topics were not developed as major research themes in the e-commerce field. Instead, they merged with some other topics.

Meanwhile, in 2011–2020 (See Fig. 5b), e-commerce research shows a diversity of themes and did not form a continued subdivision in this period. Technologies related to e-commerce maintained the field's importance in this period, and like in the first period, they were not developed as core research topics. In 2011–2020, technologies related to social commerce were popular e-commerce research topics, such as *online community*, *web 2.0*, *location-based service (LBC)*, *social media*, *virtual world*, *online review*, and *social network site*. Moreover, the new business model, *social commerce*, also developed as a research topic by 2011, and it has been widely studied in the past 10 years. We also noticed that the application of *big data* and *data analysis* has driven innovation in e-commerce and was recently has integrated with research on *AI* and *machine learning*. Notably, the *consumer* has become a hot topic in e-commerce research, and different consumer-related research themes were scattered throughout our study period—for example, *satisfaction*, *loyalty*, *trust*, *engagement*, *online review*, *personalization*, *EWOM*, *security*, and *privacy*. However, these topics were scattered throughout e-commerce research in the early 2011–2020 period but have not yet developed into core research topics.

5. Discussion and conclusion

This study identified the major research topics in e-commerce research, as well as their evolution, from 2001 to 2020 by employing co-word analysis. Our research findings provide some new insights into the e-commerce field. First, the findings in this study indicate that some research themes developed based on a cumulative tradition. For example, in the comparison of Tables 5 and 6, the topics in clusters B5 (*eGovernment*, *ICT*, *information system*, *internet*, *interoperability*) B6 (*B2C*, *marketing*, *mobile e-commerce*, *network effect*, *online payment*, *risk*, *service*), and B12 (*clustering*, *recommendation system*) dominated e-commerce research in the 2011–2020 period; however, these themes were inherited from the field's prior research in 2001–2010, such as clusters A1 (*eContract*, *eGovernment*, *interoperability*, *matchmaking*, *ontology*, *semantic web*, *service*, *XML*), A10 (*B2C*, *loyalty*, *online shopping*, *retailing*, *satisfaction*, *service quality*), and A9 (*clustering*, *recommendation system*). Therefore, these research topics have retained their continuity in e-commerce research, establishing a temporary continuity despite some topics' fading or emergence.

We also found that the themes of e-commerce research have continuously evolved over the last 20 years. New topics have always emerged and been absorbed by other research themes or replaced some old research topics in e-commerce field. For instance, B2C research was combined with different research themes over the two studied periods. In 2001–2010, the B2C topic was closely linked with the user service experience of online shopping (Cluster A10) since e-commerce was in an initial stage and consumers' reaction to e-commerce needed to be explored (See Table 5 and Fig. 5a). In the 2011–2020 period, B2C

research was combined with diverse topics, such as *marketing*, *mobile commerce*, *network effects*, *online payment*, *risk*, and *services* (Cluster B6). As B2C e-commerce matured, researchers slowly moved to the topics of *service enhancement* and *business development* (See Table 6 and Fig. 5b). B2C research has also been identified as a popular research topic in the study of. Our findings provide how B2C research has evolved over time with the development of B2C e-commerce in practice, which has not been examined in the study of Wareham et al. (2005).

Furthermore, we found that topics in e-commerce research also change with technologies' development. For example, the XML topic in Cluster A1 disappeared, and *cloud computing*, *mobile technologies*, *artificial intelligence*, and *machine learning* emerged in recent years (See Tables 5 and 6). The finding on technology as a center in e-commerce research is consistent with the findings in the studies of Ngai and Wat (2002), Shiao and Dwivedi (2013), and Yoo and Jang (2019). They also found that technology keeps as a core research theme in e-commerce research. But our findings also provide evidence on the shift of scholars' attention to innovative technology emerging over time, which has not been covered in their research. Our findings also echos to the argument from Wareham et al. (2005) that the research topics in e-commerce research should have been developed alongside new technologies continuously applied to the e-commerce domain. Therefore, future e-commerce research should not only examine traditional topics in new contexts but also explore new technologies' integration and application in the e-commerce field.

This study's results also provide useful hints about how e-commerce might develop in the near future. Specifically, the research themes in clusters B2 (*auction*, *mechanism design*, *procurement*), B4 (*privacy*, *security*), B8 (*blockchain*, *fintech*, *innovation*, *online platform*), B11 (*algorithm*, *big data*, *China*, *data analysis*, *empirical research*, *machine learning*, *web-site*), and B13 (*consumer review*; *cross-border*; *data mining*; *online review*; *sentiment analysis*) in Quadrant II (See Fig. 4b) were relatively mature and closely linked internally yet loosely related to other research themes. This finding showed a high degree of independence. Therefore, the development of research content and the direction of these topics might be somewhat limited by their failure to fit into a more central position in e-commerce research.

Almost half of the examined research subjects were clustered in Quadrant IV (See Fig. 4b), such as:

B1 (*brand*, *EWOM*, *Facebook*, *microblog*, *online community*, *social commerce*, *social influence*, *social media*, *social network*, *social network site*, *Twitter*, *web 2.0*).

B3 (*cloud computing*, *game theory*, *online advertising*, *pricing*, *retailing*, *SCM*, *search*, *strategy*).

B6 (*B2C*, *marketing*, *mobile e-commerce*, *network effect*, *online payment*, *risk*, *service*).

B7 (*consumer behavior*, *consumer experience*, *loyalty*, *online shopping*, *personalization*, *purchase behavior*, *reputation*, *satisfaction*, *SEM*, *service quality*, *sharing economy*, *survey*, *TAM*, *trust*).

B9 (*engagement*, *enjoyment*, *gamification*, *innovation diffusion*, *mobile apps*, *mobile technology*, *technology adoption*, *virtual world*).

B10 (*adoption*, *AI*, *business model*, *case study*, *CRM*, *crowdfunding*, *crowdsourcing*, *decision support*, *experiment*, *group buying*, *internet of things*, *LBS*, *mobile service*, *online banking*, *open data*, *P2P*, *product review*, *smart city*, *social capital*, *software*, *value*).

These findings suggest that the research topics in these six research clusters lie at the heart of e-commerce research. The findings on the research topics in these six clusters are consistent with some findings from prior research, such as business model and technology have also been identified as the future e-commerce research trends in the study of Yoo and Jang (2019), and social commerce, online reviews, and social media as future e-commerce research trends in the study of Mou et al. (2019). However, in this study we found that these topics are more closely related to other topics, which has not been covered in prior research. Moreover, their internal knowledge structures are not yet stable, requiring further development to allow for a high degree of

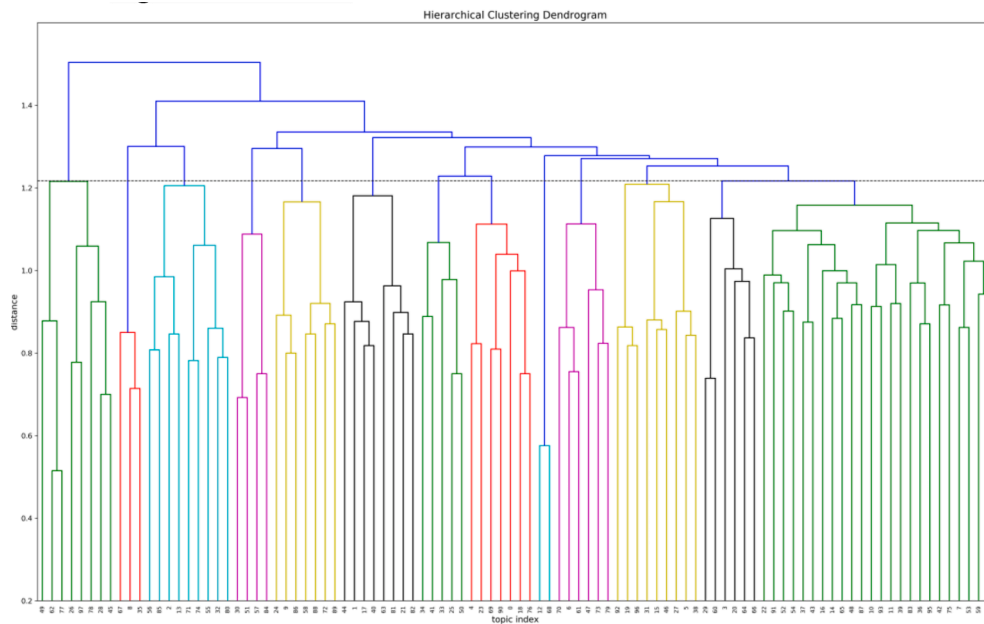


Fig. A1. Hierarchical clustering analysis of keywords in 2001–2010.

internal aggregation before they mature. Thus, our findings suggest that these topics point to the direction of e-commerce research in the coming years.

Meanwhile, the research themes in clusters B5 (*eGovernment, ICT, information system, internet, interoperability*) and B12 (*clustering, recommendation system*) were located in Quadrant I (See Fig. 4b). This position shows that B5 and B12 are the core themes in the e-commerce field. Notably, clusters A9 and B12 (See Tables 5 and 6) contain exactly the same topics, indicating that recommendation systems have continuously occupied a core position in e-commerce research over the past two decades. They differ in that, under researchers' attention, these clusters have evolved from low centrality during the first studied period to high centrality during the second studied period. Specifically, the B12 cluster—headed by *eGovernment*—closely related to Cluster A1 (*eContract, eGovernment, interoperability, matchmaking, ontology, semantic web,*

service, XML) in 2001–2010, and they were both core research themes located in Quadrant I for their respective periods (See Fig. 4a and 4b). In 2001–2010, research associated with *eGovernment* mainly attempted to solve the underlying metadata construction and interoperability problems of the *eGovernment* and *eContract* themes through *semantic web, ontology technologies, and services*. After years of e-commerce development, the research on interoperability in the *eGovernment* theme has matured. With the internet's renewal, *eGovernment* research has gradually focused more on systems' external relevance, such as ICT-related technologies. Therefore, while the sub-themes contained in these two research clusters may change as society evolves, the overall direction of *recommendation systems* and *eGovernments* that they represent are likely to continue receiving significant attention from e-commerce researchers.

Finally, new technologies have been increasingly integrated into e-

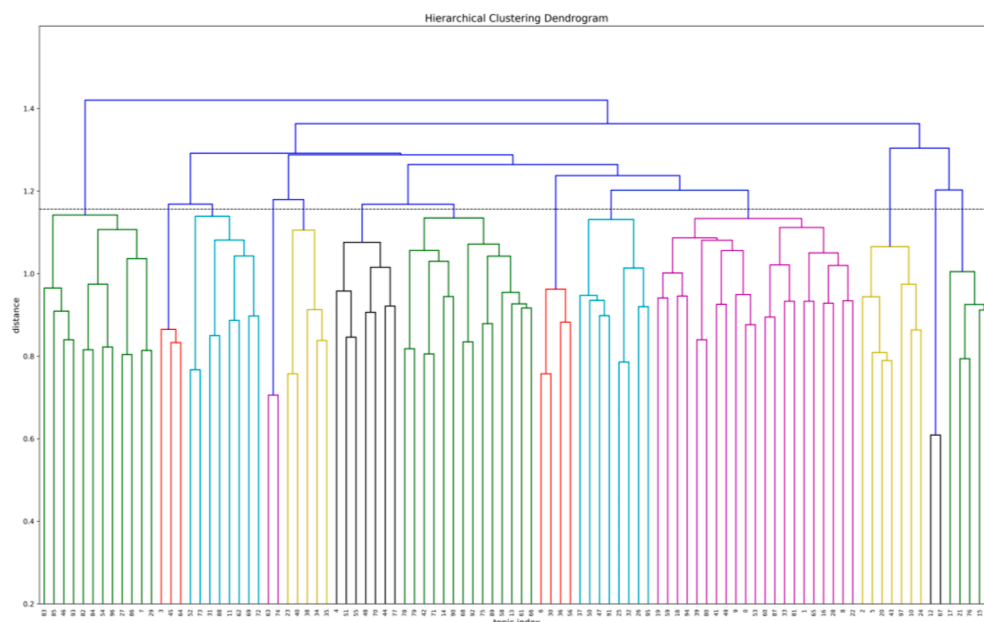


Fig. A2. Hierarchical clustering analysis for keywords in 2011–2020.

commerce to meet customers' needs, and research topics on these technologies have also been absorbed into the existing e-commerce literature to further advance the field's knowledge structure. Thus, e-commerce scholars have also exhibited greatly differing preferences for research topics. Nonetheless, their collective efforts will shape the future development of e-commerce research.

6. Limitations and avenues for future research

This study's limitations should be acknowledged. First, the current study only collected the keywords from articles published in seven leading journals in the e-commerce field from 2001 to 2020. Accordingly, our findings on the evolution of e-commerce research might not reflect broader research in the field. Therefore, future research should consider including more keyword data from more journals related to e-commerce research. Second, the current study focused on examining the knowledge structure of e-commerce research based on co-word analysis. Future research should investigate the contributions of scholars, journals, research institutions, and countries to e-commerce research via applying co-author and co-citation analysis.

CRedit authorship contribution statement

Yang Bai: Conceptualization, Writing – original draft. **Hongxiu Li:** Conceptualization, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was partly supported by the National Social Science Foundation of China under Grant 19ZDA345.

Appendix

References

- Bai, Y., Li, H., Liu, Y., 2021. Visualizing research trends and research theme evolution in E-learning field: 1999–2018. *Scientometrics* 126 (2), 1389–1414. <https://doi.org/10.1007/s11192-020-03760-7>.
- Bauin, S., Michellet, B., Schweighoffer, M.G., Vermeulin, P., 1991. Using bibliometrics in strategic analysis: "understanding chemical reactions" at the CNRS. *Scientometrics* 22 (1), 113–137. <https://doi.org/10.1007/BF02019278>.
- Beals, E.W., 1984. Bray-curtis ordination: An effective strategy for analysis of multivariate ecological data. *Adv. Ecol. Res.* 14, 1–55. [https://doi.org/10.1016/S0065-2504\(08\)60168-3](https://doi.org/10.1016/S0065-2504(08)60168-3).
- Borgatti, S.P., Everett, M.G., 1999. Models of core-periphery structures. *Social Networks* 21 (4), 375–395. <https://doi.org/10.1364/OE.19.0000B1>.
- Bray, J.R., Curtis, J.T., 1957. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecol. Monogr.* 27 (4), 325–349. <https://doi.org/10.2307/1942268>.
- Burt, R.S., 1992. *Structural holes: The social structure of competition*. Harvard University Press, pp. 8–10.
- Callon, M., Courtial, J.P., Turner, W.A., Bauin, S., 1983. From translations to problematic networks: An introduction to co-word analysis. *Soc. Sci. Inf.* 22 (2), 191–235. <https://doi.org/10.1177/053901883022002003>.
- Callon, M., Courtial, J.P., Laville, F., 1991. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics* 22 (1), 155–205. <https://doi.org/10.1007/BF02019280>.
- Cho, J., 2020. Intellectual structure evolution of open access research observed through correlation index of keyword centrality. *Scientometrics* 125 (3), 2617–2635. <https://doi.org/10.1007/s11192-020-03682-4>.
- Cui, Y., Mou, J., Liu, Y., 2018. Knowledge mapping of social commerce research: a visual analysis using CiteSpace. *Electron. Commer. Res.* 18 (4), 837–868. <https://doi.org/10.1007/s10660-018-9288-9>.
- de la Hoz-Correa, A., Muñoz-Leiva, F., Bakucz, M., 2018. Past themes and future trends in medical tourism research: A co-word analysis. *Tourism Management* 65 (4), 200–211. <https://doi.org/10.1016/j.tourman.2017.10.001>.
- Ding, Y., Chowdhury, G.G., Foo, S., 2001. Bibliometric cartography of information retrieval research by using co-word analysis. *Inf. Process. Manage.* 37 (6), 817–842. [https://doi.org/10.1016/S0306-4573\(00\)00051-0](https://doi.org/10.1016/S0306-4573(00)00051-0).
- Du, S., Li, H., 2019. The knowledge mapping of Mobile Commerce Research: A visual analysis based on I-Model. *Sustainability* 11 (6), 1–26. <https://doi.org/10.3390/su11061580>.
- Fedorko, I., Bacik, R., Gavurova, B., 2018. Technology acceptance model in e-commerce segment. *Manag. Mark.* 13 (4), 1242–1256. <https://doi.org/10.2478/mmcks-2018-0034>.
- Field, J., Clarke, K., Warwick, R., 1982. A Practical Strategy for Analysing Multispecies Distribution Patterns. *Mar. Ecol. Prog. Ser.* 8 (2), 37–52. <https://doi.org/10.3354/meps008037>.
- Fronzetti Colladon, A., 2018. The Semantic Brand Score. *Journal of Business Research* 88 (7), 150–160. <https://doi.org/10.1016/j.jbusres.2018.03.026>.
- He, Q., 1999. Knowledge Discovery Through Co-Word Analysis. *Library Trends* 48 (1), 133–159.
- Ilmudeen, A., 2021. Big Data, Artificial Intelligence, and the Internet of Things in Cross-Border E-Commerce. In: Hoque, M.R., Bashaw, R.E. (Eds.), *Cross-Border E-Commerce Marketing and Management*. IGI Global, pp. 257–272.
- Javid, E.T., Nazari, M., Ghaeli, M.R., 2019. Social media and e-commerce: A scientometric ana. *International Journal of Data and Network Science* 3 (3), 269–290. <https://doi.org/10.5267/j.ijdns.2019.2.001>.
- Jayaprakash, T.P., Wagner, E.C., Van Schalkwyk, J., Albert, A.Y.K., Hill, J.E., Money, D. M., Hemmingsen, S.M., Castillo, E., Janssen, P.A., 2016. High diversity and variability in the vaginal microbiome in women following Preterm Premature Rupture of Membranes (PPROM): A prospective cohort study. *PLoS ONE* 11 (11), 1–19. <https://doi.org/10.1371/journal.pone.0166794>.
- Kauffman, R.J., Walden, E.A., 2001. Economics and electronic commerce: Survey and directions for research. *International Journal of Electronic Commerce* 5 (4), 5–116. <https://doi.org/10.1080/10864415.2001.11044222>.
- Landauer, T.K., Dumais, S.T., 1997. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychol. Rev.* 104 (2), 211–240. <https://doi.org/10.1037/0033-295X.104.2.211>.
- Lee, S.M., Hwang, T., Kim, J., 2007. An analysis of diversity in electronic commerce research. *International Journal of Electronic Commerce* 12 (1), 31–67. <https://doi.org/10.2753/JEC1086-4415120102>.
- Leydesdorff, L., Goldstone, R.L., 2014. Interdisciplinarity at the Journal and Specialty Level: The Changing Knowledge Bases of the Journal Cognitive Science. *J. Am. Soc. Inform. Sci. Technol.* 65 (1), 164–177. <https://doi.org/10.1002/asi.22953>.
- Lin, A.J., Hsu, C.L., Chiang, C.H., 2016. Bibliometric study of Electronic Commerce Research in Information Systems & MIS Journals. *Scientometrics* 109 (3), 1455–1476. <https://doi.org/10.1007/s11192-016-2142-8>.
- Liu, Y., Li, H., Goncalves, J., Kostakos, V., Xiao, B., 2016. Fragmentation or cohesion? Visualizing the process and consequences of information system diversity, 1993–2012. *European Journal of Information Systems* 25 (6), 509–533. <https://doi.org/10.1057/ejis.2016.5>.
- Liu, L., Mei, S., 2016. Visualizing the GVC research: a co-occurrence network based bibliometric analysis. *Scientometrics* 109 (2), 953–977. <https://doi.org/10.1007/s11192-016-2100-5>.
- Mou, J., Cui, Y., Kurcz, K., 2019. Bibliometric and visualized analysis of research on major e-commerce journals using citespace. *Journal of Electronic Commerce Research* 20 (4), 219–237. <https://doi.org/10.1145/3154943.3180441>.
- Muñoz-Leiva, F., Viedma-del-Jesús, M.I., Sánchez-Fernández, J., López-Herrera, A.G., 2012. An application of co-word analysis and bibliometric maps for detecting the most highlighting themes in the consumer behaviour research from a longitudinal perspective. *Qual. Quant.* 46 (4), 1077–1095. <https://doi.org/10.1007/s11135-011-9565-3>.
- Ngai, E.W.T., Wat, F.K.T., 2002. A literature review and classification of electronic commerce research. *Information and Management* 39 (5), 415–429. [https://doi.org/10.1016/S0378-7206\(01\)00107-0](https://doi.org/10.1016/S0378-7206(01)00107-0).
- Nielsen, A.E., Thomsen, C., 2011. Sustainable development: The role of network communication. *Corp. Soc. Responsib. Environ. Manag.* 18 (1), 1–10. <https://doi.org/10.1002/csr.221>.
- Norian, J.H., Jama, A.M., Eltaieb, M.H., Ali, A., 2020. Data-Driven E-commerce Techniques and Challenges in the Era of the Fourth Industrial Revolution. *Scientific Journal of Informatics* 7 (2), 291–302. <https://doi.org/10.15294/sji.v7i2.25760>.
- Richter, A., Ng, K.T.W., Fallah, B., 2019. Bibliometric and text mining approaches to evaluate landfill design standards. *Scientometrics* 118 (3), 1027–1049. <https://doi.org/10.1007/s11192-019-03011-4>.
- Shiau, W.L., Dwivedi, Y.K., 2013. Citation and co-citation analysis to identify core and emerging knowledge in electronic commerce research. *Scientometrics* 94 (3), 1317–1337. <https://doi.org/10.1007/s11192-012-0807-5>.
- Tsai, H.H., 2015. The research trends forecasted by bibliometric methodology: a case study in e-commerce from 1996 to July 2015. *Scientometrics* 105 (2), 1079–1089. <https://doi.org/10.1007/s11192-015-1719-y>.
- Wang, C., Chen, C., 2010. Electronic Commerce Research in Latest Decade: A Literature Review. *International Journal of Electronic Commerce Studies* 1 (1), 1–14. <https://doi.org/10.7903/ijecs.898>.
- Ward, J.H., 1963. Hierarchical Grouping to Optimize an Objective Function. In *Journal of the American Statistical Association* 58 (301), 236–244. <https://doi.org/10.1080/01621459.1963.10500845>.

- Wareham, J., Zheng, J.G., Straub, D., 2005. Critical themes in electronic commerce research: A meta-analysis. *Journal of Information Technology* 20 (1), 1–19. <https://doi.org/10.1057/palgrave.jit.2000034>.
- Yoo, B., Jang, M., 2019. A bibliographic survey of business models, service relationships, and technology in electronic commerce. *Electron. Commer. Res. Appl.* 33, 1–24. <https://doi.org/10.1016/j.elerap.2018.11.005>.
- Yu, L., 2019. E-Commerce Models, Players, and Its Future. In: Khosrow-Pour, M. (Ed.), *Advanced Methodologies and Technologies in Digital Marketing and Entrepreneurship*. IGI Global, pp. 193–204.
- Zipf, G.K., 1949. *Human Behavior and the Principle of Least Effort*(pp. Addison-Wesley, p. 573).
- Zong, Q.J., Shen, H.Z., Yuan, Q.J., Hu, X.W., Hou, Z.P., Deng, S.G., 2013. Doctoral dissertations of Library and Information Science in China: A co-word analysis. *Scientometrics* 94 (2), 781–799. <https://doi.org/10.1007/s11192-012-0799-1>.