RESEARCH PAPER

Uncovering the identity of Electronic Markets research through text mining techniques

Nora Nahr[1] · Marikka Heikkilä[2]

Abstract

As an established academic journal in the e-commerce and digital platforms fields, Electronic Markets (EM) features a diverse range of topics and occupies a significant role in the information systems field. The study investigates EM's topic diversity over the time period 2009–2020 using a text mining analysis and a bibliometric analysis and identifies 28 cluster groups. The analysis reveals that the top three topics are 1) service quality, 2) blockchain and other shared trust building solutions, their impact and credibility, as well as 3) consumer buying behavior and interactions. EM's core identity lies in a balanced set of core themes that bring technological, business or human/ social perspectives to the research of networked business and digital economy. This includes research on digital and smart services, applications, consumer behavior and business models, as well as technology and e-commerce data. Ethical and sustainability related topics are however still less present in EM.

Keywords Electronic Markets journal · Core identity · Text mining analysis · Cluster analysis · Unsupervised machine learning · Bibliometric analysis

JEL Classification A10 · C38

Introduction

As an academic journal established in the field of information systems (IS) in 1991, *Electronic Markets* (EM) deals with the variety of social, economic and societal influences that information technology (IT) exerts on the interaction processes between companies and their customers. This has resulted in a growing body of topics dealing with networked business and the digital economy.[1] At the same time, the diversity of topics also increases the need for scientifically grounded knowledge to create a sound knowledge base for decision-making at the organizational and societal levels when introducing, developing and implementing goods and services. Over the past years, EM has published a significant number of articles relating to the field of the digital economy, thus addressing the rise of new technological and societal advancements in this domain. The relevance and popularity of such research is reflected in the journal's reputation, among other factors. EM holds a number of top rankings and recently achieved a high impact factor in comparison with other IS journals.[2] To underscore its increasing relevance, it is helpful to periodically evaluate this diversity of topics using rigorous analytical methods such as structured literature analyses, meta-analyses, or quantitative and qualitative content analyses. This helps to ensure continuous quality assessment and topic timeliness. An extensive analysis of a domain's or a journal's diversity is of great relevance in describing and comprehensively defining its core identity and uncovering not only established topics and thematic trends that have arisen over time, but also unexplored

---

✉ Nora Nahr
  nora.nahr@uni-passau.de
  https://www.wi.uni-
  passau.de/en/informations-it-servicemanagement/

  Marikka Heikkilä
  marikka.heikkila@utu.fi

[1] Chair of Information Systems (Information and IT Service Management), University of Passau, Innstr. 43, 94032 Passau, Germany

[2] School of Economics, University of Turku, Rehtorinpellonkatu 3, 20500 Turku, Finland

---

[1] https://www.springer.com/journal/12525/aims-and-scope

[2] http://www.electronicmarkets.org/about-em/ranking

topics that may be candidates for future research efforts (Goyal et al., 2018; Hassan & Loebbecke, 2017). This has already been proposed and implemented for related domains, such as knowledge management (Serenko, 2013) or the IS discipline in general (Jeyaraj & Zadeh, 2020). In the IS discipline, which is still a relatively young research field, the regular analysis of research streams and the diversity of topics is particularly essential for uncovering developments and trends and establishing guidelines for future research agendas (Jeyaraj & Zadeh, 2020; Palvia et al., 2015).

Such efforts have already been made for research on electronic markets as well, where multiple studies have been published in which the authors analyzed the thematized topics and their development over time by applying research methods such as automated content analyses or literature reviews (e.g., Alt & Klein, 2011; Pucihar, 2020). Similarly, in thematically related conferences such as *BLED*, one can find further examples of studies that, based on the same background, investigate conference publications thoroughly on a meta level and thus draw conclusions and suggest recommendations for action in the field of e-business (e.g., Clarke, 2012; Dreher, 2012; Fteimi et al., 2020; Pucihar, 2020).

In this context, the methods used for data collection and analysis are also undergoing a change such that both are trending toward the use of machine learning algorithms (Buchkremer et al., 2019). Evidently, there is a need for rapid and extensive modes of inquiry into existing bodies of knowledge, but at the same time, we have to consider the trustworthiness of the sources and the reputation of the publishers.

Our paper contributes to the aforementioned research efforts and seeks to answer specific calls for further research. In line with previous studies (e.g., Pucihar, 2020) employing automated techniques to analyze articles published in EM, we ultimately follow the same overarching goal of detecting the identity of the discipline. However, from a methodological viewpoint, we employ two methods. First, we use a document-based cluster analysis approach that represents unsupervised machine learning, a text mining technique that has not been used to analyze the journal's content and topics thus far. Thereafter, we apply a bibliometric approach to analyze the citations of the articles in EM. This allows us to evaluate not only the volumes, but also the impact of the various topics published in EM. One contribution of our paper is thus the application of an innovative methodological approach that combines results from clustering and bibliometric analysis. Using an analysis period from the years 2009 to 2020 further distinguishes this paper from other recent studies.

To analyze the core identity of the EM, our paper addresses the following research questions:

RQ1: Which key topics have been discussed in the EM journal over the research period and which of these topics predominate?
RQ2: Which of the key topics are most impactful within the research community?

To address the aforementioned research questions, we apply a comprehensive document-based cluster analysis as part of a text mining analysis on 330 research publications published in the EM since 2009 to 2020. As a first result, 28 topical cluster groups are derived, which provide information about frequent topics discussed in the journal's publications and their relevance. Subsequently, we evaluate as a complementary step which authors and papers were most cited in the community with a bibliometric analysis utilizing the *Scopus* and *Web of Science* (formerly *ISI)* databases. The results are reviewed in the context of the previously generated cluster groups.

The remainder of this paper is structured as follows. First, we discuss recent literature, which has contributed to the discovery of the core of knowledge within the EM community in particular and the IS domain in general. In Sect. 3 we introduce our research methods. Section 4 presents the main results of the text mining and bibliometric analyses. We proceed with a thorough discussion of our results and conclude with a summary, in which we describe several contributions and derive implications for future research.

Theoretical background

There are several articles focusing on the content analysis of major IS journals. An analysis of abstracts in *Management Information Systems Quarterly* (MISQ) using the software *Leximancer* by Indulska et al. (2012) showed how MISQ's dominant theme drifted from systems development in the 1970s to the use of information in organizations in the early 2000s. On the other hand, Sidorova et al. (2007) applied latent semantic analysis and text mining methods to investigate 1,615 abstracts (from 1985 to 2006) from MISQ, *Information Systems Research* (ISR) and the *Journal of Management Information Systems* (JMIS) with respect to emerging and declining research topics, respectively. The analysis took temporal differences and dynamics in different time spans into consideration and showed how, at the turn of the century, the attention of these three journals focused on online consumer behavior, website design and trust in IT-enabled relationships. The results were included in a research agenda and emphasized the need to pay more attention to the "rigor" factor henceforth.

The same methodological approach was applied in a more recent study by Goyal et al. (2018), who also aimed to identify research topics in the IS domain during the period from

2000 to 2017. Focusing on four top IS journals—MISQ, ISR, JMIS and *Journal of the Association for Information Systems* (JAIS)—the authors provided an overview of thematic trends first for the entire corpus and then broken down by journal. Subsequently, they performed a cluster analysis to derive eleven main research topics. The most dominant research theme by far in all four journals was *knowledge management*, which was followed by *technology adoption, e-commerce, recommender systems* and *security*. A manual content analysis of seven major IS journals (for the time period between 2004 and 2013) performed by Palvia et al. (2015) showed, in turn, that the largest research theme in general was *e-commerce*. IS adoption and usage articles were mainly published in *Information & Management* (I&M), MISQ and JAIS, whereas e-commerce articles were specifically most dominant in ISR, JMIS and I&M. These analyses showed the popularity of e-commerce research. However, their results showed only high-level categories and cannot be utilized to characterize the types of topics that are discussed within the e-commerce research theme.

EM was established in 1991, and since then, it has followed a highly interdisciplinary perspective that reflects the belief that only an interdisciplinary approach can cover the full range of possible impacts of electronic markets on economic systems, and thus on societies.[3] Alt and Klein (2011) presented an analysis of the themes and topics in electronic markets research, paying specific attention to the growing proliferation of the discipline. The authors analyzed important developments over 20 years in this field and summarized their insights in three main perspectives on electronic markets (i.e., economic environment, governance model, and business model) together with three essential drivers for these perspectives (i.e., technology push, market dynamics, and institutional design). Subsequently, six propositions were made to foster the discussion and pave the way for future trends. The authors also pointed out the need for a systematic, sustainable and interdisciplinary perspective in electronic markets.

On the occasion of EM's 30th anniversary, Pucihar (2020) analyzed the contents of 211 of its publications and 356 BLED papers via an automatic content analysis using *Leximancer* software. The analysis, which covered data from 2012 to 2019, led to the derivation of concepts (i.e., words) linked in the form of concept maps. The latter showed connections between the themes and their respective concepts, while identifying themes that were either particularly hot or irrelevant. For the EM journal in particular, four central themes were identified: *information, service, business* and *online*. The data from the BLED conference revealed a deviation to the journal, as its second most relevant

theme—*process*—did not appear. The remaining four most important themes of the conference were *business, social, information* and *health.* The author used the results to develop future research perspectives and called for further emphasis on responsible and sustainable design solutions and business models.

The preceding remarks show the diversity of analysis techniques that are applied in the IS discipline to identify the investigated research themes and topical trends.

By specifically focusing on automatically grouping articles published in EM with similar topics rather than on the temporal component, we continue this tradition using a quantitative research approach and introduce cluster analysis as a promising text mining technique to determine similar cluster groups of documents and reveal major themes in EM. A bibliometric analysis is further employed to investigate the topics and their relevance and answer both research questions. The methodological approach is described in detail in the following section.

Research design

Text mining is defined as a "*knowledge intensive process in which a user interacts with a document collection […] using a suite of analysis tools*" (Feldmann & Sanger, 2007, p. 1). By exploring the corpus of literature through the application of algorithms and methods from interdisciplinary research fields (e.g., natural language processing, information retrieval), insightful knowledge, new patterns and correlations can be discovered. As a result from this typical process flow of text mining (cf. Feldmann & Sanger, 2007), semantically related themes and document clusters, topics with related terms or simply a frequency count list can be derived. Simultaneously with text mining analysis, we also carried out traditional bibliometric analysis of the same research articles. The detailed process flow for our combined text and bibliometric analysis together with a summary of the applied analysis tasks used in this study is visualized in Fig. 1.

Overall, we performed the analysis in five main steps, during which multiple tasks took place. The research questions and the overall research aim were defined in Step 1. The collection of data was subsequently initiated in Step 2. We proceeded with the preprocessing analysis in Step 3, which consisted of multiple actions. The process continued by applying dedicated algorithms and methods to analyze the text and discover interesting patterns and insights in Step 4. In Step 5 of the study, the results were evaluated, discussed and deployed. As our research questions have been described above, we focus hereinafter on the remaining four steps.
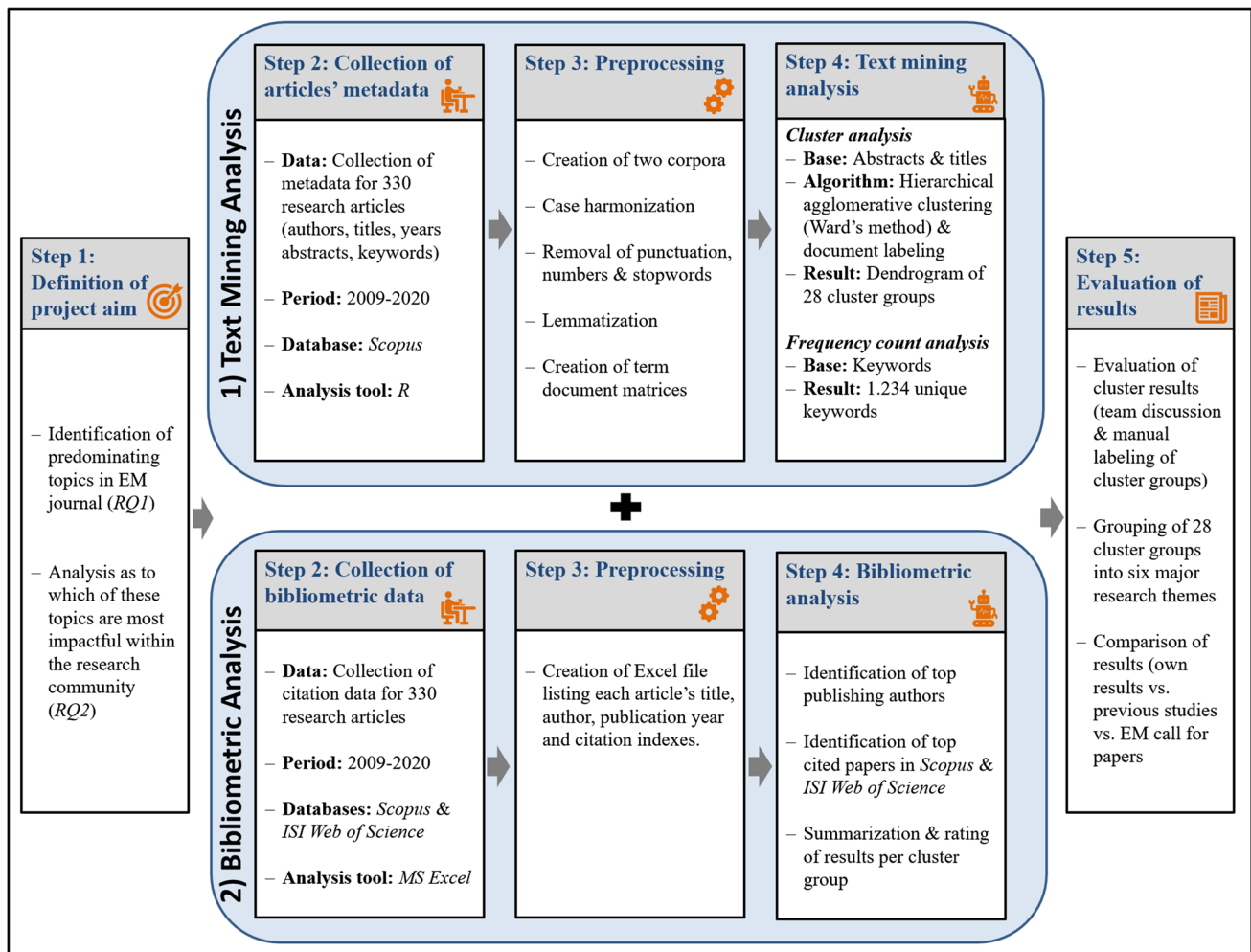

Fig. 1 Detailed research design for journal publication analysis

Data collection

In Step 2, we collected the dataset for the text mining analysis using the Scopus database, which contains as a scientific metadatabase a comprehensive collection of current and scientific publications, including EM.

The data collection covered the time span from 2009 up to 2020. Since our focus was on the analysis of research articles, we excluded editorials, discussion papers, interviews or errata. This refinement resulted in a collection of 330 datasets consisting of the titles, abstracts and keywords, which were exported to a.csv file. Since scientific abstracts generally represent a summary of the main insights of a paper, these selection criteria fit the purpose of our analysis well.[4] For the subsequent bibliometric analysis, we also recorded citation frequencies for all 330 articles via Scopus and the Web of Science databases. This allowed us to compare and contrast citation frequencies across more than one database.

Figure 2 visualizes the yearly distribution of the final dataset of EM articles and shows an increasing trend, in which the peak was reached in 2019 with 60 articles published. Since data collection was completed in August 2020, we included also 45 research articles published by August 2020.

We performed the text mining analysis using *R*, an open-source software environment for statistical data analysis (Feinerer et al., 2008), while the subsequent complementary bibliometric analysis was done in *MS Excel*.

Preprocessing

In Step 3 for text mining analysis, we divided the entire dataset into two separate corpora. Corpus 1 included all available titles and abstracts, and Corpus 2 consisted of

---

[4] The dataset used in the analysis can be requested on demand by e-mail.

Fig. 2 Yearly distribution of EM analysis data (x-axis: years; y-axis: dataset items)
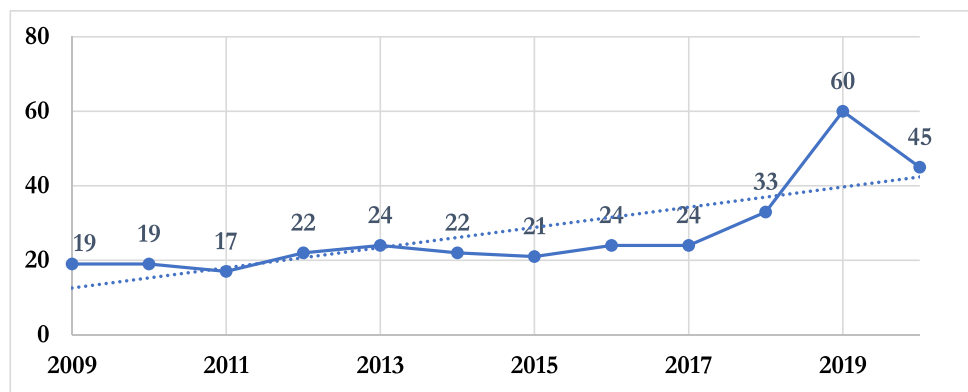
keywords for all analyzed articles. Whereas a cluster analysis was performed on all available titles and abstracts in Corpus 1, we decided to separately analyze the keywords in Corpus 2 using a frequency count analysis. This allowed us to compare single and compound keywords (which are originally separated by semicolons in the dataset) with the themes identified through the cluster analysis (which is performed on running text and therefore reveals unigrams only).

Prior to the core text mining analysis, in which the clustering algorithms were applied, among others, some preprocessing steps were necessary to harmonize both corpora.

We undertook some preprocessing steps on all running text in Corpus 1. We first performed case harmonization to ensure that different writing forms of terms (e.g., _business_ vs. _Business_) are considered as a single term. Furthermore, we also removed numbers and punctuation marks. In addition to general stopwords elimination (e.g., prepositions, articles), some names and term sequences (e.g., _University of St. Gallen_) that were mentioned at the end of the majority of the abstracts were excluded as well due to their irrelevance for the cluster analysis. Last, to group inflected term forms and harmonize their different notations to their lemma, we implemented and applied a lemmatization algorithm. A document term matrix was finally created to store information on the frequency of the relevant terms in each corpus using a mathematical representation. This matrix served as the base for the subsequent cluster analysis.

Due to their comma-separated presentation, the keywords in Corpus 2 required fewer preprocessing steps than the running text in Corpus 1. Overall, we also performed case harmonization to revert all terms to their lowercase form and to obtain consolidated results when counting keyword frequencies in later steps. The removal of some punctuation marks (e.g., dots, hyphens) as well as possible extra whitespace was also performed. The keywords were identified by semicolon separation and

saved in the form of a document term matrix. The citation data, consisting of article title, author(s), publication year and citation indexes, was collected in an Excel file. It was possible to analyze these data without any further preprocessing.

Data analysis

Cluster analysis Generally, a cluster analysis "_classifies a sample of entities into meaningful, mutually exclusive groups based on similarities among the entities_" (Balijepally et al., 2011, p. 376). As a powerful classification tool, cluster analysis is used in research, among other things, to reveal relationships between the explored data. Dependent on the data and the analysis goal, a variety of clustering algorithms can be chosen from to perform the clustering task (Balijepally et al., 2011). We performed agglomerative hierarchical clustering that has the advantage of the bottom-up and pairwise assignment of documents to clusters, which are represented in a tree-based form (i.e., dendrogram), thus enabling a detailed tracking of cluster formation (Zhao et al., 2005). During an iterative stepwise procedure, two similar or homogenous documents are assigned to a cluster based on their (dis)similarity, which is computed using an appropriate similarity measure (Huang, 2008). Once all documents are assigned to one single cluster (at the top of the dendrogram), the cluster analysis terminates. We computed the cosine similarity, which quantifies the similarity of sparse documents that usually consist of a multitude of words (attributes) using a vectorized representation. This similarity measure has wide application in text mining. With results ranging between the values 0 and 1 (1 means that a document pair is exactly alike and that its respective vectors are aligned), the similarities between document pairs are determined based on the cosine angles of their vectors. According to the following mathematical formula, their dot product is divided by the vector's magnitudes (Han et al., 2011; Li & Han, 2013; Vijaymeena & Kavitha, 2016):

$$Cos(x^{\rightarrow}, y^{\rightarrow}) = \frac{x^{\rightarrow} \bullet y^{\rightarrow}}{|x^{\rightarrow}| \, |y^{\rightarrow}|}$$

Next, we chose the clustering algorithm, a crucial decision for the further analysis (Balijepally et al., 2011). To ensure stability of results, we first computed the agglomerative coefficient based on four common hierarchical clustering algorithms (single, complete, average, ward) to calculate the strength of the cluster structure (values closer to 1 indicate a stronger structure). With a value of 0.9946, Ward's method indicated the strongest clustering structure for our dataset, leading us to select this method. By applying Ward's clustering algorithm, we also relied on a popular method in the domain of linguistic analysis (Szmrecsanyi, 2012). *Ward's* clustering (also called the minimum variance method) has the advantage of building compact, even-sized clusters (Murtagh & Legendre, 2014), which allows compensating for the drawbacks and computing requirements of other clustering methods. Due to the large dataset that was classified in the cluster analysis, we implemented a method that enabled us to derive related groups of clusters. The algorithm thus summarized the stepwise clustered documents automatically according to thematically related main cluster groups.

When determining the final amount of cluster groups, we first run several iterations with different numbers of cluster groups and then assessed visually the dendrogram and also compared both the interpretability and meaningfulness of the resulting cluster groups. Both methods originally belong to the set of methods that an analyst may invoke to determine the optimal number of clusters in cluster analyses in general, a challenge for which no one standard solution exists (Balijepally et al., 2011). Thus, the number of cluster groups that should be set for the algorithm was discussed by the research team and tested with different iterations (10, 12, 15, 18, 20, 23 25, 28, 30 and 40). Subsequently, we checked in which run the most meaningful results were achieved and, finally, chose 28 cluster groups. In this context, "meaningful" implies that the generated cluster groups are neither too generic nor too detailed (e.g., readability and expressiveness of cluster groups, no significant thematic overlaps between them) so that an observer is able to gain a compact overview of the various clustering topics without sacrificing clarity.

For each clustered article, three terms were automatically determined as labels, which sufficiently identified the respective article. For this purpose, the algorithm used the previously generated document term matrix as base. Since this matrix included information about the frequencies of terms per document, the algorithm was able to select those terms which appeared at least three times per document and as a label set.

Frequency count analysis By applying a frequency count analysis, we aimed to gain information on the frequency of unique keywords that appeared in our dataset. We analyzed unigrams (i.e., single words such as *benefit* or *success*) as well as multigrams (i.e., compound words such as *business process*). As a final step in this analysis, we summarized different lexical occurrences of a keyword (e.g., terms and their abbreviations).

Bibliometric analysis In addition to the application of various text mining algorithms, we performed bibliometric analyses to complement the main results of cluster analysis through quantitative techniques and broaden our knowledge acquired about the research field (Liu et al., 2020). We thereby adhere to the tradition, in which other EM authors conducted bibliometric analyses for different purposes in EM (e.g., Fischbach et al., 2011; Liu et al., 2020). In bibliometric analyses, researchers generally evaluate bibliographic data on records collected from a publication database and apply mathematical and statistical techniques. In doing so, it is possible to uncover interesting insights and summarize the structure of the dataset as well as the publications' content (Donthu et al., 2021; Hassan & Loebbecke, 2017; Zupic & Čater, 2015). Advantages to qualitative review-based analyses stem among others from the possibility of explorative pattern detection, an improved understanding of the research field, and the objective analysis (Daim et al., 2006; Fischbach et al., 2011; Liu et al., 2020).

By following the workflow proposed by Zupic and Čater (2015) according to five steps (cf. Fig. 1), we carried out the bibliometric analysis supplementary to the cluster analysis and based on our research questions. We utilized two widely known citation databases—Web of Science and Scopus—from which we retrieved the number of citations for each of the 330 articles and then calculate the most-cited authors and the most-cited publications. Thus, we were able to link the most-cited papers to the cluster groups determined previously by the cluster analysis. The bibliometric analysis as well as the visualization of its results were conducted in MS Excel.

Evaluation and interpretation of results

An important part of each text mining project is the evaluation and verification of the results. Depending on the type of data analyzed, this can be done, for example, using statistical methods or in a qualitative manner through team discussions and a detailed inspection of the results.

We evaluated the results of our cluster analysis in multiple ways. For example, we checked the correctness of the automatic assignment of the articles to the clusters by randomly inspecting some papers per cluster group and

checking their thematic similarity and association with the respective cluster group. We also validated the manual labeling process of the cluster groups (i.e., their naming) with a meaningful theme. One of the authors and two experienced master's students studying IS performed the initial labeling process independently after a short briefing and coded the 28 cluster groups, beginning with the most frequent terms contained therein as well as their personal judgment. Subsequently, the analysts' results were compared with each other. In the case of 14 cluster groups and their labels, all three analysts fully agreed in their selection. In the case of ten other cluster groups, the semantic meaning of the chosen label was similar, although all analysts chose different wordings. However, no immediate agreement could be reached on the labels of the four remaining cluster groups. Of these four cases, the team discussed the coding results and jointly selected a meaningful final label. As we favor descriptive labeling, we decided not to provide names that were too short to describe the contents of the cluster groups as holistically as possible. Last, to give the reader a wider overview of the clustering results and to allow easier comparison to previous research, we manually categorized the cluster groups obtained by the automated cluster analysis into six major themes.

Interpreting the results from the bibliometric analysis involved discussing key findings (cf. next section) and examining the results with regard to the cluster analysis to link them to each other.

Main results of text mining analysis

Clustering analysis—results and description

Next, we present the results of the cluster analysis. Due to the dendrogram size, we include it in the electronic appendix of the manuscript. The dendrogram illustrates the general formation of the analyzed articles into 28 cluster groups. Table 1 lists an overview of all 28 cluster groups, a selection of their prevailing terms, and their article coverage. For identification purposes, the original cluster group ID (C_ID) is shown in Column 1, whereas the numbers in Column 2 depict the cluster groups in descending order according to the articles' number. The respective major topics determined are shown in Column 3. Selected terms that occur frequently (> =twice) in the respective cluster groups are listed in the Column 4, while Column 5 indicates the number of articles per cluster group (in absolute and percentage values). The last column shows the research theme.

The three largest cluster groups are (in descending order): *service quality* (#28, core theme: digital and smart products and services), *blockchain and other shared trust building solutions, their impact and credibility* (#10, core theme:

technology) and *consumer buying behavior and interactions* (#14, core theme: consumer behavior/online interactions).

In the following, we will dissect each of these top groups to illustrate them in detail.

The largest cluster group (see Fig. 3) resulting from the analysis comprises 26 articles and addresses the topic *service quality*. The articles thereby reveal a clear tendency toward linking new technological trends and the automation of services into the context of fields such as marketing, tourism, hospitality, automotive and health care. Key terms mentioned in this context include *data, service, robot, firm, design* and *travel*. The articles in this category discuss service quality from multiple angles. For instance, the findings of Akter et al., (2010, 2013) show that service quality and perceived trust explain continuance intentions, and that that service quality has a positive effect on customer satisfaction. Sigala (2015) and Huotari and Hamari (2017), in turn, discuss gamification elements from the perspective of service quality and marketing. The implications of using robots in services —such as whether they will provide service enhancement or cost reduction (Belanche et al., 2021)— is discussed in several papers (e.g., Belanche et al., 2021; Hofmann et al., 2020; Manthiou et al., 2021). Furthermore, customer-centered services and business models are analyzed (e.g., Menschner et al., 2011; Setia et al., 2015). The cluster group thus ties in with thematic intersections of some other cluster groups we identified (e.g., #1, #20 and #27). All in all, the cluster group discusses service quality in many contexts, such as mHealth (e.g., Akter et al., 2010, 2013; Wiegard & Breitner, 2019), hospitality and travel (e.g., Manthiou et al., 2021; Novak & Schwabe, 2009; Rhee & Yang, 2015; Schmidt-Rauch & Schwabe, 2014), financial services (e.g., Chompis et al., 2014; Stamenkov & Dika, 2016; Tu, 2012) and automotive (e.g., Grieger & Ludwig, 2019).

We labeled the second-largest cluster group #10 (in terms of article coverage) as *blockchain and other shared trust building solutions, their impact and credibility* (see Fig. 4). It covers 25 articles (8%). Popular terms in this cluster group are *blockchain, technological, marketplace, benefit, competition, risk, cost* and *trust.* Earlier papers typically discuss risks and trustworthiness in online business (e.g., Gao & Liu, 2014; Keating et al., 2009; Munnukka & Järvi, 2014; Nicolaou, 2011; Sha, 2009) and the more recent ones thematize adoption of blockchain technologies and their impact on trust and organization and business (e.g., Albrecht et al., 2020; Derks et al., 2018; Kollmann et al., 2020; Marella et al., 2020; Ostern, 2020; Weking et al., 2020a, b). In addition to the introduction of technologies (e.g., Ostern, 2020) and Bitcoin cryptocurrency (e.g., Derks et al., 2018), aspects of management and the attainment of competitive advantages (e.g., Shao et al., 2012) and network effects play an important role (e.g., Bons et al., 2020), as do the associated introduction and implementation costs. It is also worth

Table 1 Clustering analysis results of EM publications

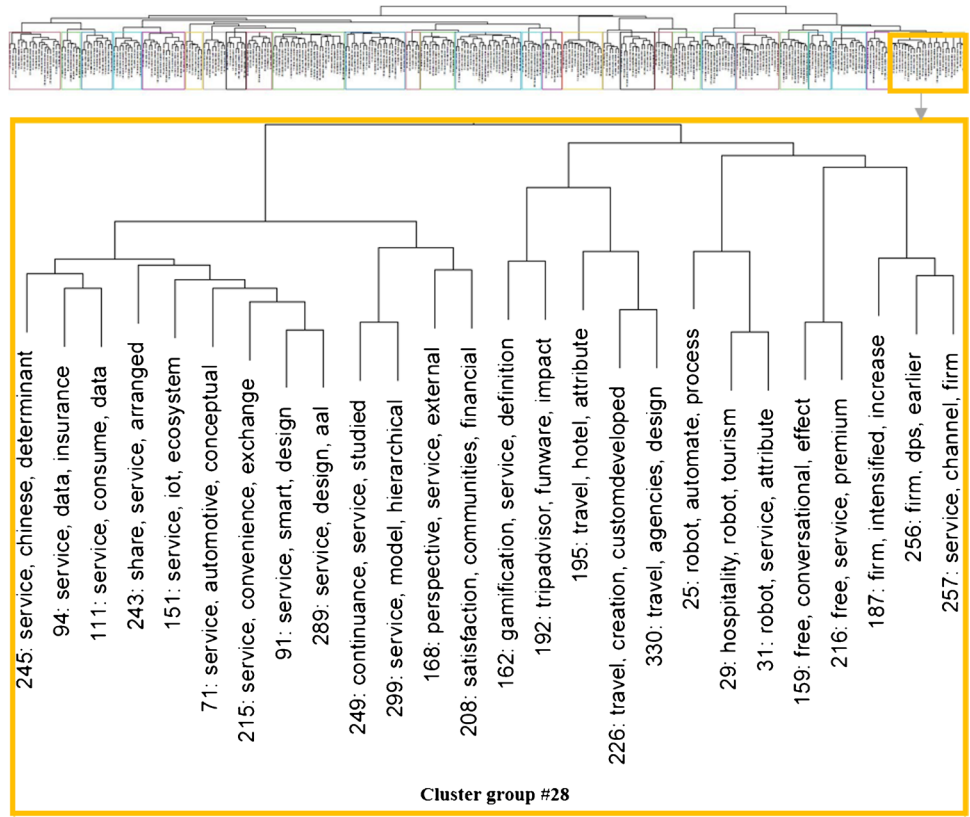| C_ID	Rank	Cluster group topic	Dominant terms per cluster group	# Articles	Core Theme
28	1	Service quality	attribute, data, design, firm, free, robot, service, travel	26 (8%)	Digital & smart products & services (17%)
24	16	Custom-developed services	custom-developed, service	10 (3%)	
7	18	Information services and firm image	information, intention	8 (2%)	
25	19	Investments in cloud services	cloud, custom-developed, investment,	8 (2%)	
27	20	Smart services	service, smart	8 (2%)	
10	2	Blockchain and other shared trust building solutions, their impact and credibility	assurance, benefit, blockchain, competition, cost, cultural, marketplace, risk, structural, technological, trust	25 (8%)	Technology (23%)
11	4	Business software development and standards	business, change, development, enterprise social software, research, standard, software, virtual	21 (6%)	
9	17	Technology adoption	adoption	9 (3%)	
2	21	Digital platforms and ecosystems	ecosystem, platform	7 (2%)	
15	23	Recommender systems	data, recommender, system	7 (2%)	
16	24	Big data analytics	commerce	7 (2%)	
14	3	Consumer buying behavior and interactions	bid, behavior, consumer, custom-developed, online, performance, player, purchase	23 (7%)	Consumer behavior/online interactions (17%)
5	6	Online social networks in business context & their governance	business, governance, network, social, peer	15 (5%)	
21	15	Auctions management	auction, bid, market	10 (3%)	
12	28	Search engines and user search behavior	search	5 (2%)	
1	5	Electronic markets applications	bank, data, driver, car, electronic, information, market, trade	18 (5%)	E-commerce applications (20%)
19	10	Product data, retail and return policies	data, effect, product, return	12 (4%)	
26	12	Mobile platform applications	health, mobile, platform, service, user	12 (4%)	
3	13	Digital innovation	digital, innovation, industrial, organization	11 (3%)	
6	25	Crowdfunding success	crowdfunding, success, factor, project	6 (2%)	
20	27	Pricing management	price, consumer	6 (2%)	
23	7	Business models	business, mature, model, process, tool, start-up	15 (5%)	Business models (11%)
22	11	Business process design	business, design, process	12 (4%)	
8	22	Supply chain management	chain, RFID, supplier	7 (2%)	
17	8	Data management	data, business, decision, market, personal, share	14 (4%)	E-commerce data (13%)
13	9	Word of mouth & consumer-centered sentiment analysis	consumer, behavior, brand, engagement, post, sentiment	12 (4%)	
4	14	Research on electronic markets & social media	academia, market, media, social, research	10 (3%)	
18	26	(Online) reviews	review, online	6 (2%)	

mentioning that in some articles, the cultural aspect and the incentives for use in a business setting are addressed and credited with the implementation of the technology.

In cluster group #14 (see Fig. 5), attention is directed toward *consumer buying behavior and interactions*. This item ranks third out of 28 cluster groups in terms of
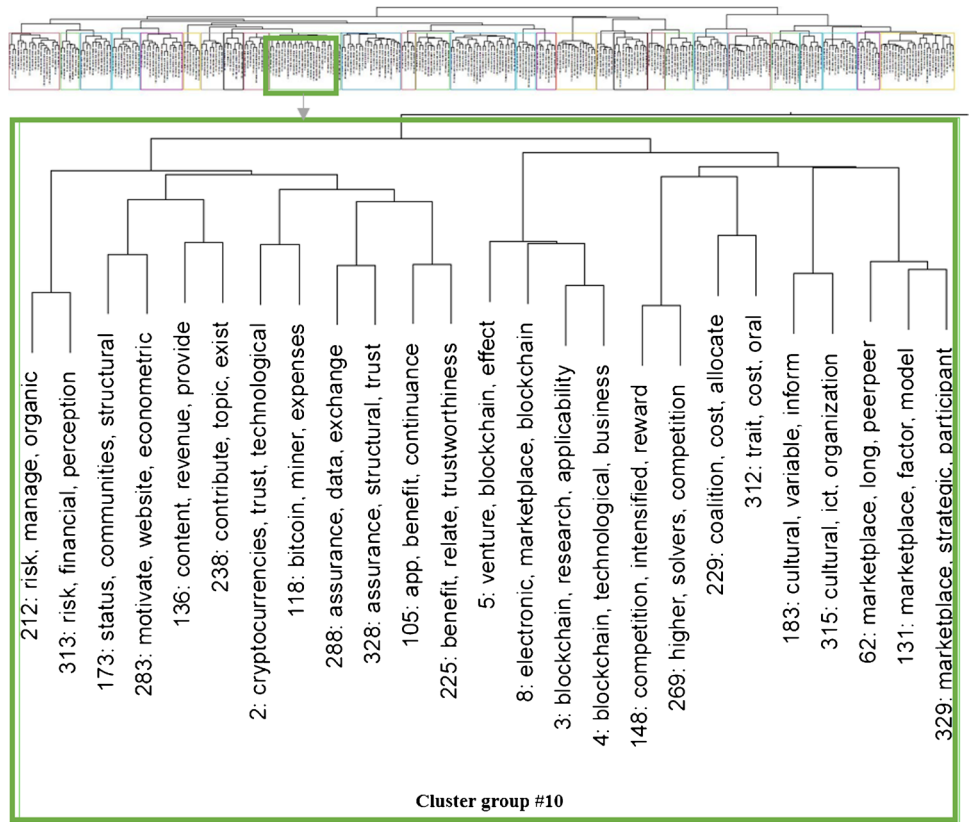
coverage. 23 (7%) of the articles are clustered together, which often mention terms such as *behavior* (e.g., Liu & Sutanto, 2012), *bid*, *purchase*, (e.g., Pappas et al., 2014), *online* and *performance* (e.g., Ghazali et al., 2016). Other interesting labels include *communities*, *emotion* and *group buying*. The papers typically discuss e.g., buyer's behavior,
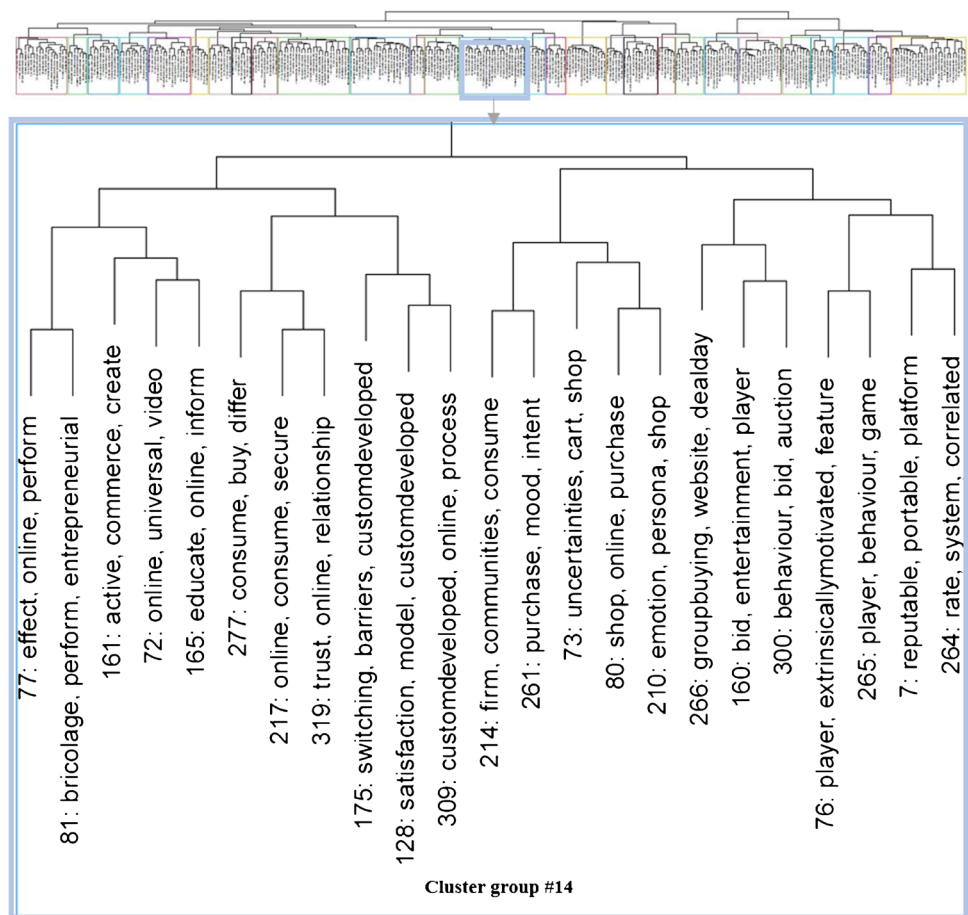
Fig. 3 Top 1 cluster group (#28)—Service quality

Cluster group #28

245: service, chinese, determinant
94: service, data, insurance
111: service, consume, data
243: share, service, arranged
151: service, iot, ecosystem
71: service, automotive, conceptual
215: service, convenience, exchange
91: service, smart, design
289: service, design, aal
249: continuance, service, studied
299: service, model, hierarchical
168: perspective, service, external
208: satisfaction, communities, financial
162: gamification, service, definition
192: tripadvisor, funware, impact
195: travel, hotel, attribute
226: travel, creation, customdeveloped
330: travel, agencies, design
25: robot, automate, process
29: hospitality, robot, tourism
31: robot, service, attribute
159: free, conversational, effect
216: free, service, premium
187: firm, intensified, increase
256: firm, dps, earlier
257: service, channel, firm

Fig. 4 Top 2 cluster group (#10)—Blockchain and other shared trust building solutions, their impact and credibility

Cluster group #10

212: risk, manage, organic
313: risk, financial, perception
173: status, communities, structural
283: motivate, website, econometric
136: content, revenue, provide
238: contribute, topic, exist
2: cryptocurrencies, trust, technological
118: bitcoin, miner, expenses
288: assurance, data, exchange
328: assurance, structural, trust
105: app, benefit, continuance
225: benefit, relate, trustworthiness
5: venture, blockchain, effect
8: electronic, marketplace, blockchain
3: blockchain, research, applicability
4: blockchain, technological, business
148: competition, intensified, reward
269: higher, solvers, competition
229: coalition, cost, allocate
312: trait, cost, oral
183: cultural, variable, inform
315: cultural, ict, organization
62: marketplace, long, peerpeer
131: marketplace, factor, model
329: marketplace, strategic, participant

Fig. 5 Top 3 cluster group (#14)—Consumer buying behavior and interactions

intentions and perceptions (e.g., Constantiou et al., 2012; Hayne et al., 2010; Hwang, 2009; Liao et al., 2010; Liu & Sutanto, 2012; Pappas et al., 2014; Penttinen et al., 2019; Riquelme & Román, 2014; San Martín & Jiménez, 2011; Zeng et al., 2019). For example, Liu and Sutanto (2012) show how the sense of urgency in auctions makes buyers purchase earlier. A further look at the titles revealed the authors' interest in modeling buying (e.g., Zeng et al., 2019). Some papers are dedicated to customers' emotions and the relevance of privacy and security during interactions with other parties (e.g., Riquelme & Román, 2014).

Keyword analysis—results and description

Next, we analyzed the keywords in Corpus 2 to evaluate whether they adequately reflect the contents of the publications. Table 2 lists, in descending frequency, all keywords that were mentioned at least five times, resulting in 25 unique keywords.

The fact that many keywords represent compound terms allows us to draw a stronger conclusion about the context of the terms. For interpretation purposes, we consolidated different spellings of keywords (e.g., keywords and their abbreviations) subsequent to the analysis.

In summary, 1,234 unique keywords occurred throughout the corpus, of which 84% were mentioned only once. The 25 keywords in Table 2 account for 12% of the overall cumulative term frequency. The top five keywords are *electronic commerce* (23 counts), *business models* (17 counts), *social media* (11 counts), *case studies* and *internet of things* (10 counts each). Other core keywords deal with financial topics (e.g., *blockchain*, *online auction*), the acceptance of digital platforms and technologies (e.g., *technology adoption*, *satisfaction*, *trust*) and cloud technologies (e.g., *cloud computing*, *crowdfunding*). It is also noteworthy that various research methods (e.g., *case study*, *literature review*, *taxonomy*) appear among the most common terms, which validates the use of the proposed research methods in EM.

Bibliometric analysis—results and description

We identified 795 unique authors in our dataset (without further differentiating between primary or coauthors). Figure 6 presents a visualization of 27 authors who published more than two papers.

The distribution of the number of authors per paper is as follows: 6 papers: 1 author; 5 papers: 5 authors; 4 papers: 10

Table 2 Keywords in EM proceedings with a frequency > = 5

Keywords	Occurrence frequency
Electronic commerce	23
Business model	17
Social media	11
Case study	10
Internet of things	10
Blockchain	9
Privacy	9
Electronic market	8
Multisided platform	8
Literature review	7
Online auction	7
Social network	7
Cloud computing	6
Crowdfunding	6
Data quality	6
Electronic word of mouth	6
Satisfaction	6
Big data	5
Design science	5
Digitalization	5
Social commerce	5
Taxonomy	5
Technology adoption	5
Trust	5
Value cocreation	5

authors; 3 papers: 11 authors; 2 papers: 59 authors; 1 paper: 709 authors (mean: 1.17).

Directing our attention to the citation frequencies per paper, we integrate data from two citation databases: Web of Science and Scopus. This allows us to cover similarities between both citation databases (see Table 3).

For each of the 330 articles in our dataset, we determined the number of citations (absolute and percentage) per database to identify the top five cited papers on both Scopus and Web of Science databases. By comparing the results, it can be seen that Akter and Wamba (2016)—who rank first on the Web of Science database with an absolute citation count of 146—and Huotari and Hamari (2017) are ranked among the top five cited papers on both the Scopus and Web of Science databases and are thematically assigned to the cluster groups labeled *big data analytics* (#16) and *service quality* (#28), respectively. The other top rankings differed between Scopus and Web of Science. In a final step, we summarized these data for each of the 28 cluster groups and generated a rating based on the mean citations per database (see Table 4). In Scopus, the top ranking is held by the cluster group entitled *smart services* (#27), as the articles in this cluster group were cited an average of 69 times. The articles belonging to the cluster group *big data analytics* (#16) were, on average, most cited in Web of Science and second most cited in Scopus.

The results show that, except for *service quality* (#28), cluster groups that were ranked largest in the clustering analysis actually do not receive the most attention from academics, as measured by the number of citations. In contrast, citation analysis highlights the importance of allowing for a variety of in topics in EM. Cluster groups such as *big data analytics* (#16) and *mobile platform applications* (#26) are receiving much interest, as they are frequently cited in both Web of Science and Scopus, but in the cluster analysis they are not similarly ranked.

Fig. 6 Top authors published in EM (n > 2 papers)
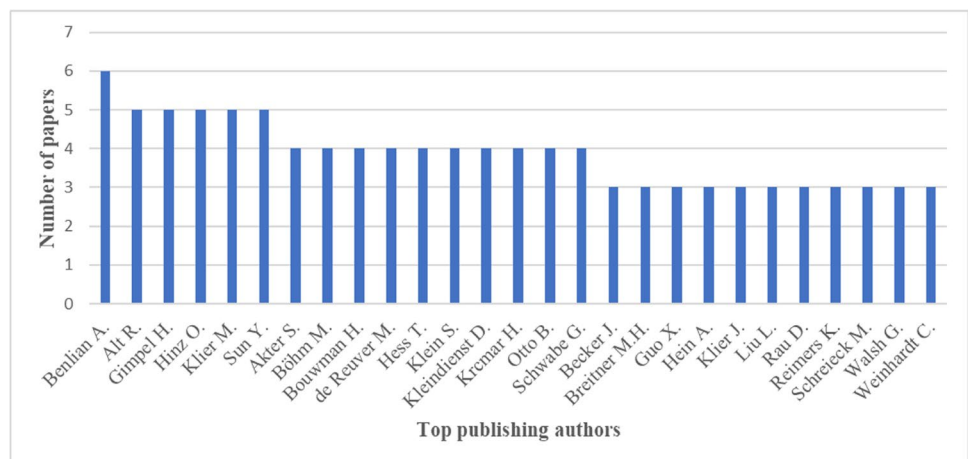
Table 3 Top five cited papers in the Web of Science and Scopus databases (authors sorted alphabetically)

Authors	Paper title	Web of Science database		Scopus database		Cluster group
		Counts	Rank	Counts	Rank	
(Akter & Wamba, 2016)	Big data analytics in E-commerce: a systematic review and agenda for future research	146 cites (9%)	(1st)	171 cites (4%)	(2nd)	Big data analytics (#16)
(Akter et al., 2013)	Continuance of mHealth services at the bottom of the pyramid: The roles of service quality and trust	61 cites (4%)	(4th)			Service quality (#28)
(Gretzel et al., 2015)	Smart tourism: foundations and developments			366 cites (8%)	(1st)	Smart services (#27)
(Huotari & Hamari, 2017)	A definition for gamification: anchoring gamification in the service marketing literature	124 cites (8%)	(2nd)	153 cites (3%)	(3rd)	Service quality (#28)
(Liu & Sutanto, 2012)	Buyers' purchasing time and herd behavior on deal-of-the-day group-buying websites	51 cites (3%)	(5th)			Consumer buying behavior and interactions (#14)
(Neuhofer et al., 2015)	Smart technologies for personalized experiences: a case study in the hospitality domain			123 cites (3%)	(5th)	Smart services (#27)
(Xu & Gupta, 2009)	The effects of privacy concerns and personal innovativeness on potential and experienced customers' adoption of location-based services			152 cites (3%)	(4th)	Technology adoption (#9)
(Zhang et al., 2011)	Business engagement on Twitter: a path analysis	66 cites (4%)	(3rd)			Word of mouth and consumer-centered sentiment analysis (#13)

Table 4 Top five cluster groups according to article volumes and citations

| Rank	Clustering analysis	Web of Science		Scopus	
	Rank according to number of articles	Top cited cluster groups	Average citations	Top cited cluster groups	Average citations
1	Service quality (#28)	Big data analytics (#16)	24,43	Smart services (#27)	68,88
2	Blockchain and other shared trust building solutions, their impact and credibility (#10)	Word of mouth & consumer-centered sentiment analysis (#13)	12,58	Big data analytics (#16)	30,43
3	Consumer buying behavior and interactions (#14)	Investments in cloud services (#25)	10,63	Technology adoption (#9)	27,22
4	Business software development and standards (#11)	Service quality (#28)	9,35	Service quality (#28)	24,19
5	Electronic markets applications (#1)	Mobile platform applications (#26)	6,83	Mobile platform applications (#26)	22,50

Discussion

Revealing the core identity of the EM community

There are several interesting findings from cluster analysis. First, the cluster analysis reveals that the three largest cluster groups called *service quality* (#28), *blockchain and other shared trust building solutions, their impact and credibility* (#10) and *consumer buying behavior and interactions* (#14) are attracting great interest in the EM community. However, this does not indicate that other topics are less important. In general, technological research core theme (reflected by six cluster groups) is strongly represented, demonstrating that addressing

Fig. 7 EM's core identity is formed from a balanced set of core themes covering technological, business and human/societal perspectives in research
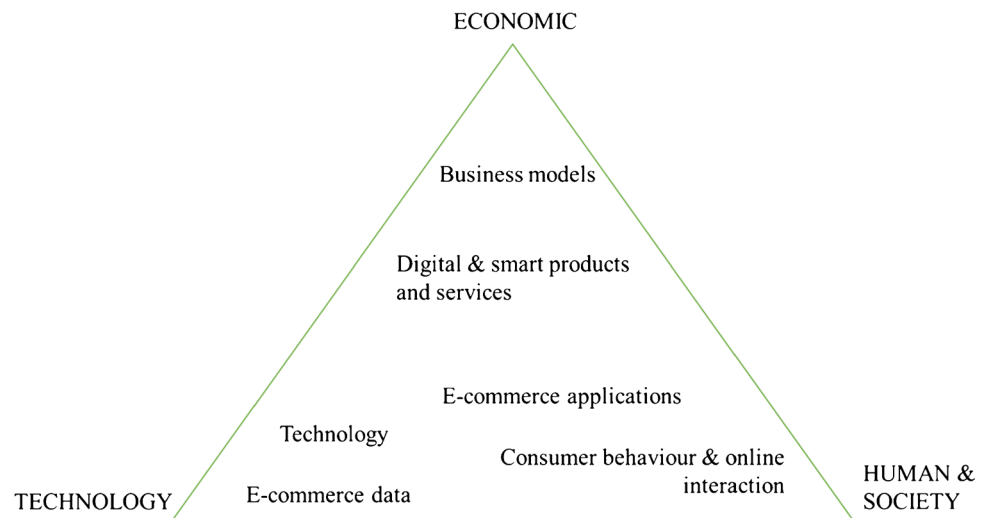
various innovative technological solutions and investigating their impact is of central importance to the community. A particularly surprising aspect of the cluster analysis is the high proportion of contributions that address the topic of blockchain and other trust enhancing technologies, as well as solution approaches in this environment. Moreover, even though artificial intelligence with its various facets, such as robotics and digital assistants was somewhat underrepresented in EM during the research period, we can observe an increasing activity in the community, especially in the last year, and forthcoming special issues are dedicated to these topics. The above reflects the focus of EM in analyses of technological innovations, and thus corresponding to the societal interest in these topics.

Another interesting finding that we draw from the cluster analysis is the observable application focus, which is particularly noticeable through six cluster groups. These articles apply IS research in diverse areas of daily life, whether in the healthcare sector, the financial sector, the automotive industry, or innovation management to name some examples. These kinds of articles are valuable in bridging the gap between research and practice and by acting as a knowledge broker between them.

To summarize the core identity of EM, we finally derived six core themes studied in EM articles:

1. Technology (23% of EM articles, six cluster groups covered), discussing adoption of technology, big data, platforms and ecosystems, blockchain and other shared trust building solutions, recommender systems and business software development and standards.
2. Applications of e-commerce (20%, six cluster groups covered), discussing electronic markets and mobile applications, pricing, crowdfunding and product data, retail and return policies.

3. Digital & smart products and services (18%, five cluster groups covered), discussing service quality, smart services, custom-developed services, cloud services investments and information services.
4. Consumer behavior and online interactions (16%, four cluster groups covered), discussing buying behavior and interactions, online social networks and auctions management, search engines and user search behavior.
5. E-commerce data (13%, four cluster groups covered) discussing data management, the application of text analysis techniques like sentiment analysis to discover EM topics as well as reviews and research in EM academia.
6. Business models (11%, three cluster groups covered), discussing business models and process designs, and supply chain management.

This diversity supports Alt and Klein's (2011) notion that electronic market research arises from IT research and economic research. Actually, we observe that EM has three main viewpoints on electronic markets research, namely *technology*, *economic*, and *human/society* (cf. Fig. 7).

Based on this conclusion, we characterize EM as a journal in the domain of networked business and the digital economy, which publishes a balanced set of articles taking technological, business and human/societal research perspectives into account.

The bibliometric analysis sheds light on the impact of the EM research. There is a varying dominance of articles and their citation frequency within the scientific databases investigated (Scopus and Web of Science). This is also reflected in Table 4, where cluster groups with different average citation counts are represented. While some cluster groups appear in both citation databases within the top five cluster groups, other cluster groups (e.g., #25, #13, #9) are dominantly represented in only one database, showing the different citation dynamics within the

databases examined. All in all, measured in average number of citations, the six core research themes (cf. Fig. 7) reach approximately similar levels, except for digital & smart products and services core theme, which clearly has received more attention in the research community. This is largely explained by the fact that several of the top cited articles belong to this core theme (Akter et al., 2013; Gretzel et al., 2015; Huotari & Hamari, 2017; Neuhofer et al., 2015).

Topic coverage by call for paper initiatives

The diversity of EM articles and the topics addressed therein can simultaneously provide valuable implications for the formulation of call for paper (CfP) initiatives. Looking at past CfPs on EM's website, we found thematic overlaps with the results of our analysis and also some differences. For example, the core theme of smart products and services which was found to have the highest impact factor in EM, was addressed by a recent CfP published in 2021 (cf. *special issue call on smart cities—smart governance models for future cities*). The same observation can be made regarding the publication of articles on innovation-related topics (cf. cluster group #3) or topics dealing with governance issues in the realm of social networks (cf. cluster group #5). This leads us to the conclusion that many submissions are made in response to CfPs and that both tasks—the publication of CfPs by the editors and the submission of articles by authors—are aligned in many cases. On the other hand, we notice that some other dominant clustering topics (e.g., cluster group #10: *blockchain and other shared trust building solutions, their impact and credibility*) as well as less dominant topics (e.g., cluster group #13: *word of mouth and consumer-centered sentiment analysis*) have not been emphasized in recent CfPs, which might deliver interesting insights for future CfPs. It should also be noted that EM has recently issued calls for submissions in the field of artificial intelligence (e.g., *special issue calls on the dark sides of AI*) and thus endeavors to address highly trending topics in this area and to link them to the topics of digital platforms and markets.

A comparative analysis of keyword and clustering themes

Keyword analysis reveals that the community places great emphasis on the use of sophisticated research methods to provide methodologically sound and validated results. An increased use of novel methodological approaches (e.g., recent methods of data analysis such as deep learning approaches) and the combination of quantitative and qualitative methodological approaches could provide valuable new insights in this field for the future.

A comparison of the list of keywords in Table 2 with those listed on the journal's website[5] reveals some commonly used terms (e.g., *electronic commerce, business model, multisided platforms, adoption, trust, survey* and *taxonomy*). Minor deviations can be observed with regard to the occurrence or dominance of some terms, such as *small and medium-sized enterprises* (no single mention in the keyword analysis) or *supply chain management* (only one single mention in the keyword analysis). This could be a possible incentive for the increased use of text mining analysis to align topics.

From a methodological perspective, this analysis shows that authors tend to use combined terms to describe their work, which permits useful deductions from both the linguistic and methodological standpoints. Text mining analyses can bring the context of the studies into the foreground by focusing on such combined terms, e.g., by using domain dictionaries (Boritz et al., 2013), as they are already applied in related disciplines such as knowledge management or the IS domain in general (Fteimi et al., 2019; Kaufmann & Bathen, 2014).

Looking at the similarities between the results of the most frequent keywords and the cluster analysis, which included the abstracts and titles, it becomes visible that the topics *electronic commerce, blockchain, online auctions* and *technology adoption* are frequently addressed in both corpora (including abstracts and keywords); thus, a comparable degree of coverage exists. Interestingly, digital or smart services are not included among the top keywords.

The use of meaningful keywords to adequately describe the content of contributions is an important prerequisite for authors in the EM community to ensure the visibility of their research activities. At the same time, this allows appropriate search results to be returned when scientific databases are queried by other researchers and interested parties.

EM topical dominance—a comparison of present and past results

Our findings complement those of Pucihar (2020) with regard to the research topics covered in EM. In Pucihar's analysis, the articles resulted in a total of 13 different themes. Of these, the four largest themes were *service*, *information*, *business* and *online*. *Service* corresponds well with the *digital & smart products and services* core theme in our results forming 17% of all EM articles. In turn, Pucihar's *information* theme is somewhat similar to our *e-commerce data* core theme covering 13% of all analyzed articles. Similarly, Pucihar's business theme corresponds with our *business models* core theme (11%). There were also some differences between the analyses. The fourth most important

---

Table 5 Results comparison and themes mapping

Core themes in this research	C_ID	Cluster group topic in this research	Themes by Pucihar (2020)
Digital & smart products & services (17%)	28	Service quality	Service
	24	Custom-developed services	
	7	Information services and firm image	
	25	Investments in cloud services	
	27	Smart services	
Technology (23%)	10	Blockchain and other shared trust building solutions, their impact and credibility	Platforms Design
	11	Business software development and standards	System Internet
	9	Technology adoption	
	2	Digital platforms and ecosystems	
	15	Recommender systems	
	16	Big data analytics	
Consumer behavior/online interactions (17%)	14	Consumer buying behavior and interactions	Customer Online Social
	5	Online social networks in business context & their governance	
	21	Auctions management	
	12	Search engines and user search behavior	
E-commerce applications (20%)	1	Electronic markets applications	Markets Electronic Mobile
	19	Product data, retail and return policies	
	26	Mobile platform applications	
	3	Digital innovation	
	6	Crowdfunding success	
	20	Pricing management	
Business models (11%)	23	Business models	Business models
	22	Business process design	
	8	Supply chain management	
E-commerce data (13%)	17	Data management	Information
	13	Word of mouth & consumer-centered sentiment analysis	
	4	Research on electronic markets & social media	
	18	(Online) reviews	

concept in Pucihar's study; *online,* only partly corresponds with our *consumer behavior/online interactions* core theme (17%). Furthermore, the biggest core theme in our results is *technology* (23%), In Pucihar's paper, these topics are scattered under several different concepts.

Table 5 illustrates how our study complements the findings of Pucihar (2020) with a more in-depth mapping and description of the topics in both studies.

Evaluating our findings against Goyal et al. (2018) and Palvia et al. (2015), it is interesting to note that the topic of *knowledge management,* which was one of the most dominant topics across IS journals, is mentioned only peripherally in the EM cluster group *business models* (#23). On the other hand, the other dominating topics in IS journals (e.g., recommender systems, e-commerce, technology adoption and security) are also present in EM. On the whole, the 28 cluster groups highlight the many facets of EM.

Last, regarding the call for more research on ethical, responsible and sustainable development of electronic markets (Alt & Klein, 2011; Pucihar, 2020), our analysis confirms that this kind of research is still largely missing, except for recent articles that discuss trust in the context of cryptocurrencies (e.g., Marella et al., 2020)) or data markets (e.g., Bauer et al., 2020) and sustainability in relation to business models (e.g., Bouwman et al., 2018; Gimpel et al., 2020).

Methodological diversity of text analysis techniques

From a methodological point of view, we can observe the following similarities and differences to Pucihar's (2020) approach.

In both studies, the data analysis is based on metadata (titles, abstracts, and keywords) of research articles. While we focused on the analysis of 330 EM articles, Pucihar performed her data analysis on a total of 567 items from

EM (211) and the BLED conference (356). However, we performed the implementation in R and used a tree-based dendrogram and tables for the visualization of the bibliometric results, while Pucihar relied on Leximancer to graphically represent the results in the form of concept maps. She employed an automatic topic modeling analysis and thus used a word-based classification approach to probabilistically derive core themes and concepts out of the corpus, which are connected to each other. In contrast, we applied a document-based approach and performed a hierarchical agglomerative cluster analysis, allowing us to classify each research article into a unique cluster based on thematic similarities between articles and by iteratively determining the most similar article pair(s) in each clustering step. Clustering thus focuses on uncovering the internal homogeneity (respectively the external heterogeneity) of clustered objects using the intra-cluster and inter-cluster distances (Weking et al., 2020a, b). This allows identifying groups of research articles based on their content similarity. The resulting 28 cluster groups are labeled with several characteristic terms and summarized finally according to major themes. In addition, we complemented our evaluation with a keyword analysis, a bibliometric analysis of citation data and a comparative analysis to put the results altogether in context.

Moreover, this study shows how the combination of machine learning techniques with traditional bibliometrics analysis is helpful in reviewing and synthesizing a comprehensive research field with a huge number of publications. Unsupervised learning methods provide a groundbreaking instrument for the future to perform similar analyses (across subjects as well as subject-specific). We see the advantages of these methods not only in their high automatization degree. Unlike supervised or even purely manual analysis procedures, this can lead to more objective results, a reduced error rate and time savings. The procedures also have a high maturity level, as they are used in various domains and are regularly tested.

With regard to the methods used in the other EM articles, we find a few examples of the application of scientometrics techniques (e.g., Fischbach et al., 2011). Cluster analysis has also been used, particularly in the last two years and often in combination with the development of taxonomies (Möller et al., 2022), e.g., to uncover archetypes of analytics-based services (Hunke et al., 2022) or to identify business model patterns and classify them along a taxonomy (Weking et al., 2020a, b). These studies use cluster analysis for a particular field of application. However, a holistic meta-perspective on all topics in EM, conducted using unsupervised cluster analysis, and combined with bibliometric analysis, does not exist in EM yet. We thus consider our study to be the most comprehensive and up-to-date analysis performed on a meta-perspective on all topics in the EM.

Summary

The overall objective of this study was to examine in depth the diversity of topics in EM journals from 2009 to 2020. While RQ1 addressed the dominance of different topics in journal submissions, RQ2 focused on which topics most influenced the EM research community. The general methodological framework for the study constituted a text mining analysis, or an unsupervised cluster analysis and keyword analysis using 330 records from the EM journal, followed by a bibliometric analysis to complement the results. While the titles and abstracts of the journal articles were used for the cluster analysis, the keywords listed by the authors served as a basis for comparison with the clustering topics. The bibliometric analysis additionally included citation data per article for the databases Scopus and Web of Science and thus allowed us to evaluate the popularity of the clustering topics.

Theoretical and practical contributions and implications

Our paper has a number of conclusions and implications for research and practice.

Conclusion #1- Revealing the core identity of EM: we can conclude based on our text mining and bibliometric analyses that during the past twelve years the core identity of the EM lies in a balanced set of core research themes that bring technological, business or human / social perspectives to the research of networked business and the digital economy. The core themes include research on digital and smart services, applications, consumer behavior and business models, as well as technology and e-commerce data. Together, they develop the identity of EM, which sets it apart from other IS journals. We consider this to be a great opportunity, as it not only promotes the continuous development of a scientific discipline, but also ensures its relevance, innovative power and sustainable value generation and growth.

Conclusion #2- The research topics and their dominance in EM: our results reveal how these core themes emerge from 28 cluster groups. This analysis answers to our first research question by providing more in-depth knowledge on the research topics and which of them are most dominating ones. The three largest cluster groups are *service quality*; *blockchain and other shared trust building solutions, their impact and credibility;* and *consumer buying behavior and interactions.*

Conclusion #3- #The impact of the topics within the research community: moreover, our bibliometric analysis responds to our second research question and shows

that the papers are actively cited by recognized scientific databases and thus attract great interest and reputation beyond the boundaries of IS. Based on average citations in Web of Science and Scopus, the top cluster groups are *big data analytics* (Web of Science) and respectively *smart services* (Scopus).

Based on the aforementioned conclusions, three key implications can be derived:

Implication #1- Exploration of EM research: the in-depth analysis of EM publications and their respective contents allows junior and senior researchers to get a comprehensive snapshot of the community's research interests and their history. The cluster dendrogram serves as a research map to understand what topics attracted the attention of the community and what are examples of yet unexplored topics. The bibliometric analysis results can be used to identify the most active authors contributing to popular topics, thus initiating collaborations opportunities in the sense of the EM research tradition. For example, based on the classified articles in the fifth cluster group, we were able to quickly identify papers with similar aims to those in our paper, such as the analyses by Clarke and Pucihar (2013) or Fischbach et al. (2011).

Implication #2- Establishment of a networking base between science and practice: our study also provides interesting insights for practitioners, considering that many cluster groups clearly highlight the application relevance of many studies through investigations in practical domains such as tourism, automotive or customer relationship management. Thus, practitioners not only benefit from getting a quick and graphical overview of publications on topics that might be of interest to them. Collaborative initiatives can also be established with the authors of the examined papers to discuss possibilities for validating scientific findings in the field.

Implication #3- Application of sophisticated machine learning techniques: this study demonstrates how sophisticated machine learning techniques can be creatively combined with traditional approaches of bibliometric analysis to advance our body of knowledge and gain new insights using automated and grounded techniques of text mining. We would like to encourage other researchers to continue this tradition and combine further techniques of machine learning to perform similar analysis and investigate the use of these methods to evaluate the findings.

Limitations and future research opportunities

Like any research paper, our article exhibits limitations and presents opportunities for future research, which we summarize in the following:

1) Database selection and analysis period: first, we deliberately chose to focus the text mining analysis on the articles available in Scopus database. This limited the analysis to articles from 2009 onward. Moreover, since we conducted the data collection process in August 2020, some of the topics that have appeared since then may be underrepresented in or even missing from our analysis. To evaluate this, we conducted a manual comparison of the new articles that have been published since the time of our data collection. The most recent topics published show a growing interest in the areas of artificial intelligence, robotics, digital assistants, and machine learning and its various applications (e.g., Nam et al., 2020; Yim et al., 2021), which is also the subject of recent calls for research. Since other IS journals (ISR, JMIS and I&M, in particular) and conferences also publish papers on similar topics, future research could incorporate these datasets as well and thus elaborate commonalities and differences thematically as well as methodologically.

2) Scope of citation indexes: in this research, we utilized citations to assess the impact of the EM articles in the scientific community. Even though bibliometric methods are established means to evaluate scientific publications (Donthu et al., 2021), their limitations are also well known. For example, Web of Science and Scopus citation indexes are based on a curated collection of documents and thus are sensitive to biases in the selection criteria of which publication outlets are included in the collection. To improve the validity of our results, we utilized two different indexes, Web of Science and Scopus. Indeed, the resulting lists of top five articles in Web of Science and Scopus look rather different: only two articles appear on both lists. However, it is shown that these databases have limited coverage in the areas of social sciences and humanities, literature written in languages other than English, and scholarly documents other than journal articles (Martín-Martín et al., 2018). One potential avenue for improving the evaluation would be to use other indexes, such as Google Scholar.

3) Further analysis of EM topics: building on the multiple research clusters identified in this paper, a further study could analyze the EM research articles even further and summarize their main scientific findings. This can be done by applying further text mining techniques (c.f. next bullet) or by using artificial intelligence tools. Moreover, the temporal evolution of topic and cluster content was not considered in our work. Thus, an interesting point for future research may be to consider different research foci in different eras (cf. Clarke & Pucihar, 2013).

4) Employment of different learning algorithms and further analysis techniques: methodologically, we employed cluster analysis as an unsupervised machine

learning technique that allows for subsequent replication and reliability assessment. However, the next step—naming and grouping of the cluster groups—requires human interpretation and thus still needs to be done manually. To the best of our knowledge, we have attempted to ensure the validity of our labeling process by having the results independently coded and evaluated by different persons. Future research could replicate the analysis by using other machine learning-based techniques for text analysis (e.g., topic modeling or deep learning-focused methods) to ensure comparability at the analysis level and delve deeper into the document database. A further interesting analysis form for future research involves the consideration of compound terms in order to provide deeper insights into the data and their context. Current approaches in text mining (such as cluster analyses) are by default based on the analysis of single terms (e.g., *commerce*, *digital*, *mobile*, *platform*), which can sometimes limit the interpretive power of the analysis. Depending on the publication language, a variety of compound terms (e.g., *digital commerce*, *mobile platform*) exist that may provide valuable insights during the analysis. One potential approach to deal with this issue is to use dictionaries that contain a predefined list of all (compound) terms to be considered in the analysis. However, these dictionaries often require a significant amount of manual effort to create and customize the list due to their domain specificity.

5) Development of a domain ontology for EM: at the same time, further research is required, for example, to develop a comprehensive domain ontology to capture the semantics behind the text. Creating such an ontology helps to specify and conceptualize important concepts in EM as well as their meanings and relationships to other concepts, thus enabling the reuse of knowledge structures.

References

Akter, S., & Wamba, S. F. (2016). Big data analytics in e-commerce: A systematic review and agenda for future research. *Electronic Markets, 26*(2), 173–194. https://doi.org/10.1007/s12525-016-0219-0

Akter, S., D'Ambra, J., & Ray, P. (2010). Service quality of mHealth platforms: Development and validation of a hierarchical model using PLS. *Electronic Markets, 20*(3), 209–227. https://doi.org/10.1007/s12525-010-0043-x

Akter, S., Ray, P., & D'Ambra, J. (2013). Continuance of mHealth services at the bottom of the pyramid: The roles of service quality and trust. *Electronic Markets, 23*(1), 29–47. https://doi.org/10.1007/s12525-012-0091-5

Albrecht, S., Lutz, B., & Neumann, D. (2020). The behavior of blockchain ventures on Twitter as a determinant for funding success. *Electronic Markets, 30*(2), 241–257. https://doi.org/10.1007/s12525-019-00371-w

Alt, R., & Klein, S. (2011). Twenty years of electronic markets research – Looking backwards towards the future. *Electronic Markets, 21*(1), 41–51. https://doi.org/10.1007/s12525-011-0057-z

Balijepally, V., Mangalaraj, G., & Iyengar, K. (2011). Are we wielding this hammer correctly? A reflective review of the application of cluster analysis in information systems research. *Journal of the Association for Information Systems*, *12*(5), 375–413. https://doi.org/10.17705/1jais.00266

Bauer, I., Zavolokina, L., & Schwabe, G. (2020). Is there a market for trusted car data? *Electronic Markets, 30*(2), 211–225. https://doi.org/10.1007/s12525-019-00368-5

Belanche, D., Casaló, L. V., & Flavián, C. (2021). Frontline robots in tourism and hospitality: Service enhancement or cost reduction? *Electronic Markets, 31*(3), 477–492. https://doi.org/10.1007/s12525-020-00432-5

Bons, R. W., Versendaal, J., Zavolokina, L., & Shi, W. L. (2020). Potential and limits of blockchain technology for networked businesses. *Electronic Markets, 30*(2), 189–194. https://doi.org/10.1007/s12525-020-00421-8

Boritz, J. E., Hayes, L., & Lim, J.-H. (2013). A content analysis of auditors' reports on IT internal control weaknesses: The comparative advantages of an automated approach to control weakness identification. *International Journal of Accounting Information Systems, 14*(2), 138–163. https://doi.org/10.1016/j.accinf.2011.11.002

Bouwman, H., Heikkilä, J., Heikkilä, M., Leopold, C., & Haaker, T. (2018). Achieving agility using business model stress testing. *Electronic Markets, 28*(2), 149–162. https://doi.org/10.1007/s12525-016-0243-0

Buchkremer, R., Demund, A., Ebener, S., Gampfer, F., Jägering, D., Jürgens, A., Klenke, S., Krimpmann, D., Schmank, J., Spiekermann, M., Wahlers, M., & Wiepke, M. (2019). The application of artificial intelligence technologies as a substitute for reading and to support and enhance the authoring of scientific review articles. *IEEE Access, 7*, 65263–65276. https://doi.org/10.1109/ACCESS.2019.2917719

Chompis, E., Bons, R. W. H., van den Hooff, B., Feldberg, F., & Horn, H. (2014). Satisfaction with virtual communities in B2B financial services: Social dynamics, content and technology. *Electronic Markets, 24*(3), 165–177. https://doi.org/10.1007/s12525-014-0160-z

Clarke, R. (2012). The first 25 years of the Bled eConference: themes and impacts. *Proceedings of the 25th Bled eConference - Special Issue* (pp. 12–192). BLED, Slovenia.

Clarke, R., & Pucihar, A. (2013). Electronic interaction research 1988–2012 through the lens of the Bled eConference. *Electronic Markets, 23*(4), 271–283. https://doi.org/10.1007/s12525-013-0144-4

Constantiou, I., Legarth, M. F., & Olsen, K. B. (2012). What are users' intentions towards real money trading in massively multiplayer online games? *Electronic Markets, 22*(2), 105–115. https://doi.org/10.1007/s12525-011-0076-9

Daim, T. U., Rueda, G., Martin, H., & Gerdsri, P. (2006). Forecasting emerging technologies: Use of bibliometrics and patent analysis. *Technological Forecasting and Social Change, 73*(8), 981–1012. https://doi.org/10.1016/j.techfore.2006.04.004

Derks, J., Gordijn, J., & Siegmann, A. (2018). From chaining blocks to breaking even: A study on the profitability of bitcoin mining from 2012 to 2016. *Electronic Markets, 28*(3), 321–338. https://doi.org/10.1007/s12525-018-0308-3

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

Dreher, H. (2012). Automatic semantic trend analysis of the Bled eConference: 2001–2011. *Proceedings of the 25th Bled eConference - Special Issue* (pp. 193–208). BLED, Slovenia.

Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, 25(5), 1–54. https://doi.org/10.18637/jss.v025.i05

Feldmann, R., & Sanger, J. (2007). *The text mining handbook.* Cambridge University Press. https://doi.org/10.1017/CBO9780511546914

Fischbach, K., Putzke, J., & Schoder, D. (2011). Co-authorship networks in electronic markets research. *Electronic Markets, 21*(1), 19–40. https://doi.org/10.1007/s12525-011-0051-5

Fteimi, N., Basten, D., & Lehner, F. (2019). Advancing automated content analysis in knowledge management research: The use of compound concepts. *International Journal of Knowledge Management, 15*(1), 53–68.

Fteimi, N., Heikkilä, M., & Heikkilä, J. (2020). Topical research cluster of BLED community – A text mining approach. *Proceedings of the33rd BLED eConference* (pp. 499–514).

Gao, H., & Liu. (2014). Relationship of trustworthiness and relational benefit in electronic catalog markets. *Electronic Markets*, 24(1), 67–75.https://doi.org/10.1007/s12525-013-0142-6

Ghazali, E., Nguyen, B., Mutum, D. S., & Mohd-Any, A. A. (2016). Constructing online switching barriers: Examining the effects of switching costs and alternative attractiveness on e-store loyalty in online pure-play retailers. *Electronic Markets, 26*(2), 157–171. https://doi.org/10.1007/s12525-016-0218-1

Gimpel, H., Graf-Drasch, V., Kammerer, A., Keller, M., & Zheng, X. (2020). When does it pay off to integrate sustainability in the business model? A Game-Theoretic Analysis. *Electronic Markets, 30*(4), 699–716. https://doi.org/10.1007/s12525-019-00361-y

Goyal, S., Ahuja, M., & Guan, J. (2018). Information systems research themes: A seventeen-year data-driven temporal analysis. *Communications of the Association for Information Systems, 43*(1), 404–431. https://doi.org/10.17705/1cais.04323

Gretzel, U., Sigala, M., Xiang, Z., & Koo, C. (2015). Smart tourism: Foundations and developments. *Electronic Markets, 25*(3), 179–188. https://doi.org/10.1007/s12525-015-0196-8

Grieger, M., & Ludwig, A. (2019). On the move towards customer-centric business models in the automotive industry – A conceptual reference framework of shared automotive service systems. *Electronic Markets, 29*(3), 473–500. https://doi.org/10.1007/s12525-018-0321-6

Han, J., Kamber, M., & Pei, J. (2011). Getting to know your data. In J. Han & M. Kamber (Eds.), *Data Mining: Concepts and Techniques* (3rd ed., pp. 39–81). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-381479-1.00002-2

Hassan, N. R., & Loebbecke, C. (2017). Engaging scientometrics in information systems. *Journal of Information Technology, 32*(1), 85–109. https://doi.org/10.1057/jit.2015.29

Hayne, S. C., Bugbee, B., & Wang, H. (2010). Bidder behaviours on eBay: Collectibles and commodities. *Electronic Markets, 20*(2), 95–104. https://doi.org/10.1007/s12525-010-0036-9

Hofmann, P., Samp, C., & Urbach, N. (2020). Robotic process automation. *Electronic Markets, 30*(1), 99–106. https://doi.org/10.1007/s12525-019-00365-8

Huang, A. (2008). Similarity measures for text document clustering. *Proceedings of the 6th New Zealand Computer Science Research Student Conference* (pp. 9–56). Christchurch, New Zealand.

Hunke, F., Heinz, D., & Satzger, G. (2022). Creating customer value from data: Foundations and archetypes of analytics-based services. *Electronic Markets, 32*(2), 1–19. https://doi.org/10.1007/s12525-021-00506-y

Huotari, K., & Hamari, J. (2017). A definition for gamification: Anchoring gamification in the service marketing literature. *Electronic Markets, 27*(1), 21–31. https://doi.org/10.1007/s12525-015-0212-z

Hwang, Y. (2009). The impact of uncertainty avoidance, social norms and innovativeness on trust and ease of use in electronic customer relationship management. *Electronic Markets, 19*(2–3), 89–98. https://doi.org/10.1007/s12525-009-0007-1

Indulska, M., Hovorka, D. S., & Recker, J. (2012). Quantitative approaches to content analysis: Identifying conceptual drift across publication outlets. *European Journal of Information Systems, 21*(1), 49–69. https://doi.org/10.1057/ejis.2011.37

Jeyaraj, A., & Zadeh, A. H. (2020). Evolution of information systems research: Insights from topic modeling. *Information & Management, 57*(4), 103207. https://doi.org/10.1016/j.im.2019.103207

Kaufmann, J., & Bathen, L. (2014). Themen und Trends in der Wirtschaftsinformatik – Eine Analyse unter Einsatz von Big-Data-Technologien. *Multikonferenz Wirtschaftsinformatik* (pp. 146–153).

Keating, B. W., Quazi, A. M., & Kriz, A. (2009). Financial risk and its impact on new purchasing behavior in the online retail setting. *Electronic Markets, 19*(4), 237–250. https://doi.org/10.1007/s12525-009-0021-3

Kollmann, T., Hensellek, S., de Cruppe, K., & Sirges, A. (2020). Toward a renaissance of cooperatives fostered by Blockchain on electronic marketplaces: A theory-driven case study approach. *Electronic Markets, 30*(2), 273–284. https://doi.org/10.1007/s12525-019-00369-4

Li, B., & Han, L. (2013). Distance weighted cosine similarity measure for text classification. In H. et al. Yin (Ed.), *Intelligent Data Engineering and Automated Learning – Lecture Notes in Computer Science* (LNCS 8206., pp. 611–618). Berlin, Heidelberg: Springer International Publishing. https://doi.org/10.1007/978-3-642-41278-3_74

Liao, C., Palvia, P., & Lin, H. (2010). Stage antecedents of consumer online buying behavior. *Electronic Markets, 20*(1), 53–65. https://doi.org/10.1007/s12525-010-0030-2

Liu, Y., & Sutanto, J. (2012). Buyers' purchasing time and herd behavior on deal-of-the-day group-buying websites. *Electronic Markets, 22*(2), 83–93. https://doi.org/10.1007/s12525-012-0085-3

Liu, W., Wang, Z., & Zhao, H. (2020). Comparative study of customer relationship management research from East Asia, North America and Europe: A bibliometric overview. *Electronic Markets, 30*(4), 735–757. https://doi.org/10.1007/s12525-020-00395-7

Manthiou, A., Klaus, P., Kuppelwieser, V. G., & Reeves, W. (2021). Man vs machine: Examining the three themes of service robotics in tourism and hospitality. *Electronic Markets, 31*(3), 511–527. https://doi.org/10.1007/s12525-020-00434-3

Marella, V., Upreti, B., Merikivi, J., & Tuunainen, V. K. (2020). Understanding the creation of trust in cryptocurrencies: The case of bitcoin. *Electronic Markets, 30*(2), 259–271. https://doi.org/10.1007/s12525-019-00392-5

Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & López-Cózar, E. D. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics, 12*(4), 1160–1177. https://doi.org/10.1007/s11192-020-03690-4

Menschner, P., Prinz, A., Koene, P., Köbler, F., Altmann, M., Krcmar, H., & Leimeister, J. M. (2011). Reaching into patients' homes – Participatory designed AAL services: The case of a patient-centered nutrition tracking service. *Electronic Markets, 21*(1), 63–76. https://doi.org/10.1007/s12525-011-0050-6

Möller, F., Stachon, M., Azkan, C., Schoormann, T., & Otto, B. (2022). Designing business model taxonomies – Synthesis and guidance from information systems research. *Electronic Markets, 32*(2), 1–26. https://doi.org/10.1007/s12525-021-00507-x

Munnukka, J., & Järvi, P. (2014). Perceived risks and risk management of social media in an organizational context. *Electronic Markets, 24*(3), 219–229. https://doi.org/10.1007/s12525-013-0138-2

Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *Journal of Classification, 31*(3), 274–295. https://doi.org/10.1007/s00357-014-9161-z

Nam, K., Dutt, C. S., Chathoth, P., Daghfous, A., & Khan, M. S. (2020). The adoption of artificial intelligence and robotics in the hotel industry: Prospects and challenges. *Electronic Markets, 31*(3), 1–21. https://doi.org/10.1007/s12525-020-00442-3

Neuhofer, B., Buhalis, D., & Ladkin, A. (2015). Smart technologies for personalized experiences: A case study in the hospitality domain. *Electronic Markets, 25*(3), 243–254. https://doi.org/10.1007/s12525-015-0182-1

Nicolaou, A. I. (2011). Supply of data assurance in electronic exchanges and user evaluation of risk and performance outcomes. *Electronic Markets, 21*(2), 113–127. https://doi.org/10.1007/s12525-011-0061-3

Novak, J., & Schwabe, G. (2009). Designing for reintermediation in the brick-and-mortar world: Towards the travel agency of the future. *Electronic Markets, 19*(1), 15–29. https://doi.org/10.1007/s12525-009-0003-5

Ostern, N. (2020). Blockchain in the IS research discipline: A discussion of terminology and concepts. *Electronic Markets, 30*(2), 195–210. https://doi.org/10.1007/s12525-019-00387-2

Palvia, P., Daneshvar Kakhki, M., Ghoshal, T., Uppala, V., & Wang, W. (2015). Methodological and topic trends in information systems research: A meta-analysis of IS journals. *Communications of the Association for Information Systems, 37*(1), 630–650. https://doi.org/10.17705/1cais.03730

Pappas, I. O., Kourouthanassis, P. E., Giannakos, M. N., & Chrissikopoulos, V. (2014). Shiny happy people buying: The role of emotions on personalized e-shopping. *Electronic Markets, 24*(3), 193–206. https://doi.org/10.1007/s12525-014-0153-y

Penttinen, E., Halme, M., Malo, P., Saarinen, T., & Vilén, V.-M. (2019). Playing for fun or for profit: How extrinsically-motivated and intrinsically-motivated players make the choice between competing dual-purposed gaming platforms. *Electronic Markets, 29*(3), 337–358. https://doi.org/10.1007/s12525-018-0298-1

Pucihar, A. (2020). The digital transformation journey: Content analysis of Electronic Markets articles and Bled eConference proceedings from 2012 to 2019. *Electronic Markets, 30*(1), 29–37. https://doi.org/10.1007/s12525-020-00406-7

Rhee, H. T., & Yang, S. (2015). How does hotel attribute importance vary among different travelers? An exploratory case study based on a conjoint analysis. *Electronic Markets, 25*(3), 211–226. https://doi.org/10.1007/s12525-014-0161-y

Riquelme, I., & Román, S. (2014). Is the influence of privacy and security on online trust the same for all type of consumers? *Electronic Markets, 24*(2), 135–149. https://doi.org/10.1007/s12525-013-0145-3

San Martín, S., & Jiménez, N. H. (2011). Online buying perceptions in Spain: Can gender make a difference? *Electronic Markets, 21*(4), 267–281. https://doi.org/10.1007/s12525-011-0074-y

Schmidt-Rauch, S., & Schwabe, G. (2014). Designing for mobile value co-creation – The case of travel counselling. *Electronic Markets, 24*(1), 5–17. https://doi.org/10.1007/s12525-013-0124-8

Serenko, A. (2013). Meta-analysis of scientometric research of knowledge management: Discovering the identity of the discipline. *Journal of Knowledge Management, 17*(5), 773–812. https://doi.org/10.1108/jkm-05-2013-0166

Setia, P., Richardson, V., & Smith, R. J. (2015). Business value of partner's IT intensity: Value co-creation and appropriation between customers and suppliers. *Electronic Markets, 25*(4), 283–298. https://doi.org/10.1007/s12525-015-0189-7

Sha, W. (2009). Types of structural assurance and their relationships with trusting intentions in business-to-consumer e-commerce. *Electronic Markets, 19*(1), 43–54. https://doi.org/10.1007/s12525-008-0001-z

Shao, B., Shi, L., Xu, B., & Liu, L. (2012). Factors affecting participation of solvers in crowdsourcing: An empirical study from China. *Electronic Markets, 22*(2), 73–82. https://doi.org/10.1007/s12525-012-0093-3

Sidorova, A., Evangelopoulos, N., & Ramakrishnan, T. (2007). Diversity in IS research: An exploratory study using latent semantics. *Proceedings of the International Conference on Information Systems* (pp. 1–19).

Sigala, M. (2015). The application and impact of gamification funware on trip planning and experiences: The case of TripAdvisor's funware. *Electronic Markets, 25*(3), 189–209. https://doi.org/10.1007/s12525-014-0179-1

Stamenkov, G., & Dika, Z. (2016). Bank employees' internal and external perspectives on e-service quality, satisfaction and loyalty. *Electronic Markets, 26*(3), 291–309. https://doi.org/10.1007/s12525-016-0221-6

Szmrecsanyi, B. (2012). *Grammatical variation in british english dialects: A study in corpus-based dialectometry*. Cambridge University Press. https://doi.org/10.1017/CBO9780511763380

Tu, H. (2012). Performance implications of internet channels in financial services: A comprehensive perspective. *Electronic Markets, 22*(4), 243–254. https://doi.org/10.1007/s12525-012-0108-0

Vijaymeena, M. K., & Kavitha, K. (2016). A survey on similarity measures in text mining. *Machine Learning and Applications: An International Journal, 3*(2), 19–28. https://doi.org/10.5121/mlaij.2016.3103

Weking, J., Hein, A., Böhm, M., & Krcmar, H. (2020a). A hierarchical taxonomy of business model patterns. *Electronic Markets, 30*(3), 447–468. https://doi.org/10.1007/s12525-018-0322-5

Weking, J., Mandalenakis, M., Hein, A., Hermes, S., Böhm, M., & Krcmar, H. (2020b). The impact of blockchain technology on business models – A taxonomy and archetypal patterns. *Electronic Markets, 30*(2), 285–305. https://doi.org/10.1007/s12525-019-00386-3

Wiegard, R. B., & Breitner, M. H. (2019). Smart services in healthcare: A risk-benefit-analysis of pay-as-you-live services from customer perspective in Germany. *Electronic Markets, 29*(1), 107–123. https://doi.org/10.1007/s12525-017-0274-1

Xu, H., & Gupta, S. (2009). The effects of privacy concerns and personal innovativeness on potential and experienced customers' adoption of location-based services. *Electronic Markets, 19*(2–3), 137–149. https://doi.org/10.1007/s12525-009-0012-4

Yim, D., Malefyt, T., & Khuntia, J. (2021). Is a picture worth a thousand views? Measuring the effects of travel photos on user engagement using deep learning algorithms. *Electronic Markets, 31*(3), 619–637. https://doi.org/10.1007/s12525-021-00472-5

Zeng, M., Cao, H., Chen, M., & Li, Y. (2019). User behaviour modeling, recommendations, and purchase prediction during shopping festivals. *Electronic Markets, 29*(2), 263–274. https://doi.org/10.1007/s12525-018-0311-8

Zhang, M., Jansen, B. J., & Chowdhury, A. (2011). Business engagement on Twitter: A path analysis. *Electronic Markets, 21*(3), 161–175. https://doi.org/10.1007/s12525-011-0065-z

Zhao, Y., Karypis, G., & Fayyad, U. (2005). Hierarchical clustering algorithms for document datasets. *Data Mining and Knowledge Discovery, 10*(2), 141–168. https://doi.org/10.1007/s10618-005-0361-3

Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods, 18*(3), 429–472. https://doi.org/10.1177/1094428114562629