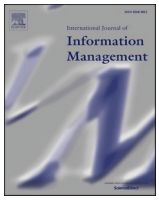
International Journal of Information Management 58 (2021) 102316

Contents lists available at ScienceDirect

International Journal of Information Management

journal homepage: www.elsevier.com/locate/ijinfomgt



Research Article

Fifty years of information management research: A conceptual structure analysis using structural topic modeling

Anuj Sharma [a,*], Nripendra P. Rana [b], Robin Nunkoo [c,d,e,f]

[a] Chandragupt Institute of Management Patna (CIMP), Patna, 800001, Bihar, India
[b] School of Management, University of Bradford, Richmond Road, Bradford, BD7 1DP, UK
[c] Department of Management, University of Mauritius, Reduit, Mauritius
[d] School of Tourism and Hospitality, University of Johannesburg, South Africa
[e] Griffith Institute for Tourism, Griffith University, Gold Coast, Australia
[f] Copenhagen Business School, Denmark

ARTICLE INFO

Keywords:
Information management
Structural topic models
Topic modeling
Generative models
Text analytics

ABSTRACT

Information management is the management of organizational processes, technologies, and people which collectively create, acquire, integrate, organize, process, store, disseminate, access, and dispose of the information. Information management is a vast, multi-disciplinary domain that syndicates various subdomains and perfectly intermingles with other domains. This study aims to provide a comprehensive overview of the information management domain from 1970 to 2019. Drawing upon the methodology from statistical text analysis research, this study summarizes the evolution of knowledge in this domain by examining the publication trends as per authors, institutions, countries, etc. Further, this study proposes a probabilistic generative model based on structural topic modeling to understand and extract the latent themes from the research articles related to information management. Furthermore, this study graphically visualizes the variations in the topic prevalences over the period of 1970 to 2019. The results highlight that the most common themes are data management, knowledge management, environmental management, project management, service management, and mobile and web management. The findings also identify themes such as knowledge management, environmental management, project management, and social communication as academic hotspots for future research.

1. Introduction

The early research on information management (hereafter, IM) reports that it embraces a variety of organizational activities starting from acquiring and integrating the information from various sources to organizing, structuring, and processing the information and finally, disseminating the information to the right person at the right time in an optimal way (Adelman & Kemp, 1970; Beimesch, 1982; Blazewicz, 1983; Brussaard, 1983, 1988; Carter, 1983; Espejo, 1983; Hoard, 1989; King & Maryanski, 1983; Lavallee, 1983; Narimatsu, 1983; Pollak, 1983; Roberts & Clarke, 1989). A few researchers also report that the IM domain is originated from database management, record management, and data processing management (Trauth, 1989). Unsurprisingly, perhaps, most of the early studies in the late-1970s and early-1980s focused on technical aspects of IM focusing on hardware and software development. During that period, the term information resources management (IRM) was associated with the management of information as well as information technologies, which are used for acquiring, storing, processing, and utilizing the information (Cheng, 1987; McClure & Hill, 1982; McClure, 1981; Otten, 1984; Roberts & Wilson, 1987). Subsequently, IM became an intuitively appealing domain to study, research and practice where the information was no more just a 'processed data' but it was regarded as an 'organizational resource' requiring effective management as other resources like man, machine, material, and money (Tranfield, 1986; Wilson, 1986). While the IRM concept was predominantly about the management of data to meet an organizational need, it could not cover the broad perspective and objective of information management, which were all about helping people and organizations to access, organize, process and utilize information efficiently and effectively.

However, the IM domain quickly observed a process of convergence and integration when researchers and practitioners started realizing that

* Corresponding author.
E-mail addresses: f09anujs@iimidr.ac.in (A. Sharma), nranap@gmail.com (N.P. Rana), r.nunkoo@uom.ac.mu (R. Nunkoo).

https://doi.org/10.1016/j.ijinfomgt.2021.102316
Received 10 July 2020; Received in revised form 10 January 2021; Accepted 10 January 2021
Available online 22 January 2021

along with technical developments in data processing, and the management and organization of people and processes are also crucial (Anderton, 1986; Ellis, 1986; Jellis, 1988; Lewis & Martin, 1989). This resulted in the separation of roles and responsibilities of a data administrator and an information manager. Still, at the same time, these domain experts started to collaborate to solve complex business information problems (Wiggins, 1988), which facilitated the convergence of a range of skills (Herring, 1991). Hence, IM became a strategic issue in organizations (Silk, 1988), a key factor in improving corporate performance (Best, 1988), an important element for encouraging competition and efficiency (Fuller-Love & Cooper, 1996), an instrument of gaining competitive advantage and a tool for reducing uncertainty (Davenport & Cronin, 1988) because information management domain provided all the theoretical basis to effectively use information in business and government organizations (Jackson, 1987). Strategic information management not only enhanced differentiation in value chain activities (Huotari, 1995), but also had a broader impact at organizational, sectoral, and societal levels (Mutch, 1996).

The exponential growth in the volume of information produced by the organizations further amplified the importance of the information management domain, which required new skills, knowledge, qualifications, and experience (Entsua-Mensah, 1996; Willard & Mychalyn, 1998) for managing information at four different levels such as information retrieval, information systems, information contexts, and information environments (Rowley, 1998). Although IM has various definitions, meanings, and interpretations, its scope is well documented by Detlor (2010), who defines IM as "the management of the processes and systems that create, acquire, organize, store, distribute, and use the information and the goal of information management is to help people and organizations access, process and use information efficiently and effectively" (p.1). Three major perspectives such as organizational, library, and personal are proposed to further define 'information management' domain that "concerns the control over how information is created, acquired, organized, stored, distributed, and used as a means of promoting, efficient and effective information access, processing, and use by people and organizations" (Detlor, 2010, p.1).

In the last five decades, IM has witnessed exponential growth covering versatile themes and topics of research. The domain knowledge and the available research literature are quite rich and deep. However, very limited efforts were made in the past to integrate this fragmented research. Hence, in the absence of such research endeavors, the overall intellectual structure of the information management research consists of islands of knowledge, which are sporadic in nature. Understanding the evolution of the prevailing intellectual structure of the IM research published in various journals and conference proceedings in the last five decades may add a significant value to the existing body of knowledge, and such study will also be potentially interesting for academics and practitioners in this field. Analyzing and describing the development of the IM domain for the last fifty years is a much-needed research endeavor, which surprisingly has not been undertaken so far. The existing retrospections on IM and related fields have focused only on specific journals (Abedin, Jafarzadeh, & Olszak, 2020; Bate & Gedam, 2019; Donthu, Kumar, Pandey, & Gupta, 2021; Dwivedi & Mustafee, 2010; Paz, Merigó, Powell, Ramaprasad, & Syn, 2020), specific duration (González-Valiente, León Santos, Arencibia-Jorge, Noyons, & Costas, 2019; Liu, Tian, Kong, Lee, & Xia, 2019), or provided an overview of its particular aspect like big data (Ahmad, Ahmed, Shah, & Ahmed, 2020; Ardito, Scuotto, Del Giudice, & Petruzzelli, 2019; Khanra, Dhir, & Mäntymäki, 2020; Parlina, Ramli, & Murfi, 2020). Hence, a study covering the evolution of the intellectual and conceptual structure of IM over a large time interval can add significant value to the body of the literature on the information management field.

Considering the unfilled gap in the literature, the current study aims to present a bird's eye view of the evolution of IM research over the last fifty years starting from 1970 to 2019 using descriptive analysis and text analytics based methodology, which is based on topic modeling technique. Bibliometrics-based descriptive analysis is a very common approach for unifying the voluminous and fragmented research into a comprehensive, systematic, objective, and transparent elaboration of the domain knowledge. The bibliometric overview in this study presents a retrospective analysis by examining the publication trend as per authors, institutions, and countries, keywords, key themes, and the methodology illustrated in Fig. 1 is consistent with the previous bibliometric methodologies proposed in the earlier works. Finally, the study performs topic modeling to visualize the latent topics and their prevalences over time, which brings novelty and differentiate the current research from the existing retrospections of the field.

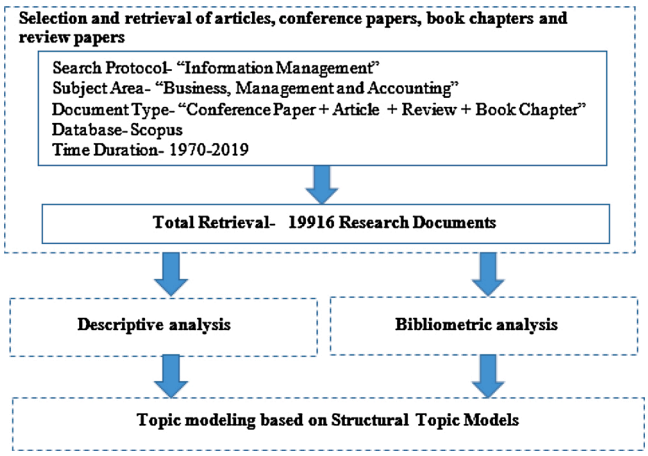


Fig. 1. Overview of the study design.

This study will serve as a gateway to the entire IM domain for readers and aspiring researchers as well as offer a comprehensive snapshot of the domain to its potential contributors. The study offers several contributions to the IM domain by summarizing the important publications and the most influential authors and visualizing the evolution of latent themes and topics over time. Apart from using the traditional methodology for research synthesis and integration like bibliometrics, this study also uses a novel methodology based on the structural topic modeling (STM) technique for extracting and visualizing the latent topics and their evolution over time. The novel STM approach discussed in this study digs deeper into the research literature and extracts the latent topics in an automatic data-driven way.

The remainder of the study proceeds as follows: Section 2 presents methodology and data. Section 3 presents the study results containing key summaries of the data analysis results. Further, Section 4 presents the discussion of the bibliometric analysis and topic modelling related findings. Finally, Section 5 concludes the study.

2. Methods and data

2.1. Bibliometric analysis

Bibliometric analysis has emerged as one of the most common, reliable, and widely accepted techniques for analyzing and summarizing voluminous and fragmented research. Bibliometric analysis originally emerged from the library and information sciences domain, where researchers started incorporating statistical quantitative techniques to transform the large volume of extant literature into a comprehensive retrospective overview (Broadus, 1987). However, many studies also support the view that bibliometric analysis gained early popularity in science and the humanities domains (Price, 1976), which was later applied to other domains like social science (Glänzel, 1996), international business (Fetscherin, Voss, & Gugler, 2010), international management (Acedo & Casillas, 2005), public policy (Sprott & Miyazaki, 2002), marketing (Kim, Kang, & Lee, 2019; Martínez-López, Merigó, Valenzuela-Fernández, & Nicolás, 2018; Martínez-López, Merigó,

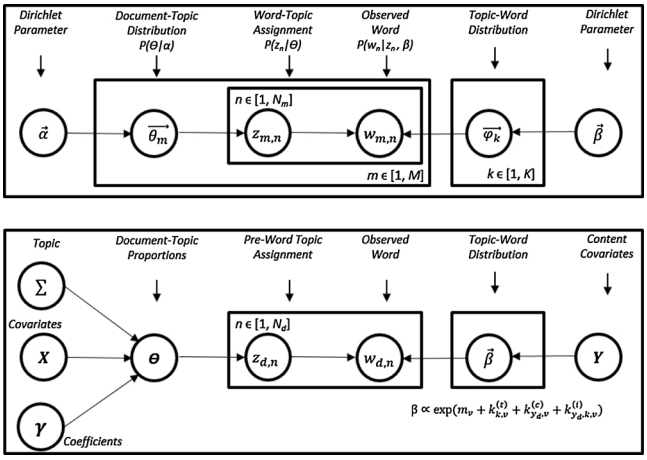


Fig. 2. Comparison of the plate notations of LDA (Adapted from source: Blei et al., 2003) and STM (Adapted from source: Roberts et al., 2014).

Gázquez-Abad, & Ruiz-Real, 2020; Valenzuela, Merigó, Johnston, Nicolas, & Jaramillo, 2017), advertising (Kim & McMillan, 2008), psychology (Tur-Porcar, Mas-Tur, Merigó, Roig-Tierno, & Watt, 2018), travel and tourism marketing (Mulet-Forteza, Martorell-Cunill, Merigó, Genovart-Balaguer, & Mauleon-Mendez, 2018), computer integrated manufacturing (Laengle, Modak, Merigó, & Sotta, 2018), communications (Pasadeos, Renfro, & Hanily, 1999), and information system (Alsudairi & Dwivedi, 2010; Avison, Dwivedi, Fitzgerald, & Powell, 2008; Dwivedi & Kuljis, 2008; Dwivedi, Lal, Mustafee, & Williams, 2008). The similarity and association among research papers, contributors, and journals can be measured using citation analysis, usually performed by assessing bibliographic coupling and co-citation analysis, generally at contributors and journal levels.

For a particular research domain, bibliometric analysis can be used to analyze the general thematic areas of the domain as a whole as well as graphically visualize the conceptual subdomains (López-Robles, Otegi-Olaso, Porto Gómez, & Cobo, 2019; Rana, Weerakkody, Dwivedi, & Piercy, 2014). For a particular journal, bibliometric analysis can be applied to understand the thematic evolution of the journal, visualize citation patterns, discern the progressive themes and envisage the avenues for future research (Cobo, Martínez, Gutiérrez-Salcedo, Fujita, & Herrera-Viedma, 2015; Donthu, Kumar, & Pattnaik, 2020; Laengle et al., 2017; Martínez-López et al., 2018; Martorell Cunill, Socias Salvá, Otero Gonzalez, & Mulet-Forteza, 2019; Singh, Dhir, Das, & Sharma, 2020).

2.2. Topic modeling based on the structural topic models

Topic modeling is one of the most useful techniques in natural language processing and text analytics which is used to analyze the text documents by extracting the underlying topics (latent themes) from the documents. This is an unsupervised machine learning technique that automatically learns and discovers the latent themes and their prevalences across a collection of documents. These discovered themes by topic modeling are part of the latent structure of text documents, and each theme consists of semantically similar words. Hence, the latent semantic structure of the text documents in a corpus can be explored by extracting the key topics from the text. The most popular topic modeling techniques are Latent Semantic Analysis (LSA; Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990; Lochbaum & Streeter, 1989), Probabilistic Latent Semantic Analysis (pLSA; Hofmann, 1999) and Latent Dirichlet Allocation (LDA; Blei, Ng, & Jordan, 2003). Inspired by the seminal research work by Blei et al. (2003), the research on LDA based topic modeling has observed tremendous growth. LDA (also called as, generative aspect model) is basically a Bayesian probability based generative probabilistic model. All LDA based topic models (Blei & Lafferty, 2005; Blei, 2012) assume that each document in a text corpus is a "*distribution over a fixed vocabulary*" (statistical mixture hypothesis) and each latent topic is a "*distribution over words with an assigned probability from 0 to 1*" (distributional hypothesis) (Blei et al., 2003, p.996). Further, the order of the documents in the corpus and the order of the words in each document is insignificant.

Structural topic modeling (STM) is a recent and sophisticated probabilistic topic modeling technique, which estimates a topic model through a fast variational approximation using an expectation-maximization algorithm (Roberts, Stewart, & Tingley, 2019). STM allows researchers to explore a document corpus with arbitrary metadata for extracting topics and estimating the topic's relationship to document metadata. A comparison of the plate notations of LDA (Blei et al., 2003) and STM (Roberts et al., 2014) is represented in Fig. 2.

Moreover, apart from extracting topics using a quantitatively methodical variational approximation model, STM can also be used to model the interaction between covariates associated with the metadata. For example, the interaction between time and extracted topics can be modeled to understand how one variable may 'moderate' the effect of another variable. Being a recent technique, STM is not widely used across disciplines, and only a few recent studies have used it for analyzing public opinions (Tvinnereim & Fløttum, 2015), survey responses (Roberts et al., 2014), news reports (Roberts, Stewart, & Airoldi, 2016) and Twitter feeds (Lucas et al., 2015).

Similar to LDA, STM is also a probabilistic generative model that defines a document generated as a mixture of hidden topics with different prevalences. Generative models explain that a particular document comes into existence by going through a data generating process using multiple topics. These K latent topics form a latent semantic space in which each word may be represented as a K-dimensional vector. Further, each topic is composed of a mixture of unigram words or bigrams terms with different prevalences, which means that each word has a probability of belonging to a particular topic. In this case, for a document, the aggregate topic proportions across all topics are one, and for a specific topic, the aggregate word probabilities of all words are also one. Finally, these latent topics can be summarized using the top-n words ranked as per the corresponding conditional probability scores called prevalences.

The full description of the convergence of structural topic modeling algorithm is out of the scope of this study. However, a simplified pseudo-code of the topic modeling process for a document corpus with associated metadata is given below, which is adapted from recent seminal work using STM for analyzing news reports (Roberts et al., 2016).

Initialize-Document corpus (indexed by d) with metadata and having vector x_d as document-level covariates, number of topics= K, total n words in a sampled document d represented as $\{w_{d,n}\}$, vocabulary size= V.

Step-1. Estimate the topic prevalence parameter (proportion vector) $\boldsymbol{\theta_d}$ for each word in a document d using a logistic-normal generalized linear model. $\boldsymbol{\theta_d}$ is generated for each word from the vocabulary of size V having the probability of $k = 1, \ldots, K$ different topics.

$$\boldsymbol{\theta}_d \sim \text{LogisticNormal}_{k-1}(\boldsymbol{\Gamma}' \boldsymbol{x_d}, \boldsymbol{\Sigma})$$

where, the topic prevalence model's coefficients are represented by $\boldsymbol{\Gamma} = [\boldsymbol{\gamma}_1 | \ldots | \boldsymbol{\gamma}_K]$ and $\boldsymbol{\Sigma}$ is a hyper-parameter modeled as a $(K-1)$ by $(K-1)$ covariance matrix.

Step-2. Generate the topical content model $\boldsymbol{\beta}$, which represents the words as a probabilistic mixture of each topic (k) using the following equation where m is the baseline word distribution vector of length V, $k^{(t)}_{k,v}$ is the topic (t) specific deviation, $k^{(c)}_{y_d,v}$ is the covariate (c) group deviation and $k^{(i)}_{y_d,k,v}$ is the interaction (i) coefficient. This study includes publication year as covariates.

$$\boldsymbol{\beta}_{d,k,v} = \frac{\exp(m_v + k^{(t)}_{k,v} + k^{(c)}_{y_d,v} + k^{(i)}_{y_d,k,v})}{\sum_v \exp(m_v + k^{(t)}_{k,v} + k^{(c)}_{y_d,v} + k^{(i)}_{y_d,k,v})}$$

Step-3. For each word in the document, ($n \in \{1,. . ., N_d\}$), the core language model (Roberts et al., 2014) represented by following unsupervised models can be used to sample a topic from a multinomial distribution over the topic prevalence parameter and for a given topic, a word is sampled using the applicable multinomial distribution.

$$\mathbf{z}_{d,n} \sim \text{Multinomial}_K(\theta_d), \text{ for n } = 1 \ldots N_d$$

$$\mathbf{w}_{d,n} \sim \text{Multinomial}_V(\mathbf{B}, \mathbf{z}_{d,n}), \text{ for n } = 1 \ldots N_d$$

Here, **B** is a global parameter, which is directly related to document-level covariates. The computation of topic prevalence parameter $\boldsymbol{\theta_d}$ in STM using a Logistic-Normal distribution differentiates it from the basic LDA (Blei et al., 2003). Due to the presence of the document metadata, STM assumes that topical prevalence (a probabilistic measure of the association between document and topic) and topical content (words of a topic) are related to some document-level covariates existent in the document metadata (Roberts et al., 2019).

In order to extract the key topics from the research documents of the IM domain and model their corresponding prevalences, the text from the article's title, keywords, abstract, etc., and the publication year were combined to create the text corpus for STM. STM based topic modeling has been found to be robust even if the words are mentioned several times at different locations in the same document (Vanhala et al., 2020).

The basic text data preprocessing involves the removal of stop words, numbers, non-English words, special characters, and punctuations. Hence, consequent to the basic cleaning of the text data prepared from the research documents published during 1970–2019, this study further processed the text corpus to make it appropriate for STM. Each document was having frequent words related to copyright information and the publishers, which were removed after a careful examination. The most frequent bigram terms were generated from the text corpus by coding a bigram tokenizer in R environment. These bigrams were then compared with the author's specified keywords. These bigram keywords were found to be useful for topic modeling. Hence, the most frequent bigram terms are concatenated in this study by which the bigram 'social network' was concatenated to 'socialnetwork' and 'big data' was converted to 'bigdata'. This study has experimented with a different number of topics (K) starting from 5 to 20, and the optimal value of K was selected based on averaged held-out likelihood from the model. To assess the topic model's quality, this study has used semantic coherence, which is a representation of the semantic association of top-n words in each topic to the documents that are highly associated with that specific topic. To make the topic modeling results reproducible, this study has used spectral initialization of the topic modeling function (Roberts et al., 2019).

2.3. Data

The bibliographic data of IM domain related articles, conference papers, book chapters, and reviews published from 1970 to 2019 was accessed and retrieved from the Scopus database in May 2020. The comprehensive coverage of multi-disciplinary research makes Scopus the most preferred destination for researchers from all domains. Previous studies have confirmed that Scopus is extensively used as a data source in many contemporary studies (Donthu et al., 2020; Singh et al., 2020). As shown in Fig. 1, the search query was designed by performing a logical combination of 'Information Management', subject area as 'Business, Management, and Accounting', and year 2020 'exclude'. Among the several other subject areas, Business, Management, and Accounting was the most prominent subject area covering a total of 20,057 documents. The search query result contained 141 undefined documents for which no detailed information was available on Scopus.

Table 1
Research documents of the IM domain.

S. No.	Document Type	Total Number
1	Article	4,735
2	Book Chapter	234
3	Conference Paper	14,560
4	Review	387
	Total	**19,916**

Therefore, a total of 19,916 research documents were used in this study after confirming the proper eradication of discrepancies, if any. Table 1 provides a snapshot of the research documents used in this study.

This study has used R environment for all the analyses, including a bibliometric overview, topic modeling, and results visualization. R is an open-source environment for statistical computations. All the modeling experiments were done on Windows 10 based computer having 16 GB RAM and 64-bit architecture.

3. Results

3.1. Most cited papers

The number of total citations indicates the overall influence of the articles on other studies (Tsay, 2009). Table 2 presents a list of the top 10 most cited studies related to the IM domain between 1970 and 2019. The most cited paper with 5410 citations was published by Alavi and Leidner (2001) in the domain of knowledge management and knowledge management systems (Alavi & Leidner, 2001). This work presents a consolidated description and definition of complex and multi-faceted concepts related to knowledge management and knowledge management systems. Further, a study exploring the strategies for managing knowledge (Hansen, Nohria, & Tierney, 1999) has attracted a total of 2523 citations, making it ranked the second in the list of most cited papers. In total, seven articles have more than 1000 citations with a significant amount of citations every year. The most recent article in the top 50 most cited papers is a manifesto created by the IEEE Task Force on 'Process Mining' in 2012 (van der Aalst et al., 2012). The next recent article in the top 50 is on 'Design science research methodology for information systems' (Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007). These details demonstrate that it takes a significant time for an article to get a considerable number of citations (i.e. more than one thousand).

3.2. Noteworthy influential authors

Table 3 illustrates a list of the most influential authors by the total number of citations and average citations per research document. Professor Dorothy Leidner from Baylor University tops the list who is followed by Professor Maryam Alavi from Scheller College of Business at Georgia Tech. Professor Varun Grover from Walton School of Business, University of Arkansas, comes next as per the total number of citations. As per the average citation per document, Professor Maryam Alavi leads the list with 5410 citations.

3.3. The most influential institutions and countries

The IM domain has attracted scholars and researchers from all over the world. The top research institutions and universities having the highest number of publications are Hong Kong Polytechnic University (total publications: 144) followed by the National University of Singapore (total publications: 81) and South China University of Technology (total publications: 78). Fig. 3 depicts the top 10 most affiliated institutions with influential authors. Among the top contributing institutions, the institutions from China have a significant share. Other contributing institutions and universities are from countries that include

Table 2
Top 10 most cited articles during 1970-2019.

Rank	Authors	Title	Year	TC
1	Alavi M., Leidner D.E.	Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues	2001	5410
2	Hansen M.T., Nohria N., Tierney T.	What's your strategy for managing knowledge?	1999	2523
3	Peffers K., Tuunanen T., Rothenberger M. A., Chatterjee S.	A design science research methodology for information systems research	2007	2333
4	Davenport T.H.	Putting the enterprise into the enterprise system.	1998	2102
5	Van Der Heijden H.	User acceptance of hedonic information systems	2004	1862
6	Sambamurthy V., Bharadwaj A., Grover V.	Shaping agility through digital options: Reconceptualizing the role of information technology in contemporary firms	2003	1644
7	Gregor S.	The nature of theory in Information Systems	2006	1578
8	Okoli C., Pawlowski S. D.	The Delphi method as a research tool: An example, design considerations and applications	2004	1524
9	Rockart J.F.	Chief executives define their own data needs.	1979	1435
10	Malhotra N.K., Kim S. S., Patil A.	Common method variance in IS research: A comparison of alternative approaches and a reanalysis of past research	2006	1393
11	Lee H.L., So K.C., Tang C.S.	Value of information sharing in a two-level supply chain	2000	1325
12	Gefen D., Straub D.W.	Gender differences in the perception and use of e-mail: An extension to the technology acceptance model	1997	1229
13	Danneels E.	The dynamics of product innovation and firm competences	2002	1132
14	King W.R., He J.	A meta-analysis of the technology acceptance model	2006	1112
15	Smith H.J., Milberg S. J., Burke S.J.	Information privacy: Measuring individuals' concerns about organizational practices	1996	1090
16	Tippins M.J., Sohi R.S.	IT competency and firm performance: Is organizational learning a missing link?	2003	946
17	Hartwick J., Barki H.	Explaining the role of user participation in information system use	1994	879
18	Payne A., Frow P.	A strategic framework for customer relationship management	2005	871
19	Walls J.G., Widmeyer G.R., El Sawy O.A.	Building an information system design theory for vigilant EIS	1992	865
20	Wixom B.H., Watson H.J.	An empirical investigation of the factors affecting data warehousing success	2001	836
21	Fui-Hoon Nah F., Lee-Shang Lau J., Kuang J.	Critical factors for successful implementation of enterprise systems	2001	835
22	Eppler M.J., Mengis J.	The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines	2004	765
23	Rosenkopf L., Almeida P.	Overcoming local search through alliances and mobility	2003	746
24	Venkatesh V., Brown S.A.	A longitudinal investigation of personal computers in homes: Adoption determinants and emerging challenges	2001	734
25	Schepers J., Wetzels M.	A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects	2007	714

Table 3
Top 10 most influential authors.

Rank	Author	Documents	Citations	Average Citations/ Document
1	Leidner D.E.	2	5825	2913
2	Alavi M.	1	5410	5410
3	Grover V.	8	2593	324
4	Sambamurthy V.	9	2564	285
5	Hansen M.T.	1	2523	2523
6	Nohria N.	1	2523	2523
7	Tierney T.	1	2523	2523
8	Rothenberger M. A.	2	2352	1176
9	Chatterjee S.	2	2345	1173
10	Tuunanen T.	2	2335	1168
11	Peffers K.	1	2333	2333
12	Straub D.W.	5	2165	433
13	Davenport T.H.	1	2102	2102
14	El Sawy O.A.	4	1888	472
15	Van Der Heijden H.	1	1862	1862
16	King W.R.	8	1765	221
17	Lee H.L.	2	1678	839
18	Bharadwaj A.	2	1655	828
19	Okoli C.	2	1602	801
20	Gregor S.	1	1578	1578
21	Pawlowski S.D.	1	1524	1524
22	Kim S.S.	2	1394	697
23	Malhotra N.K.	1	1393	1393
24	Patil A.	1	1393	1393
25	Wetherbe J.C.	5	1371	274

United States, Romania, Russian Federation, Czech Republic, United Kingdom, Malaysia, Australia, Germany, and Indonesia. This shows and confirms internationality in publication.

Table 4 highlights the top ten countries with the highest number of publications and their publications' evolution. The top countries are China, United States, Romania, Russian Federation, Czech Republic, United Kingdom, Malaysia, Australia, Germany, and Indonesia.

3.4. Top research keywords

An analysis of the research article's most common keywords discovers major areas of interest (Callon, Courtial, Turner, & Bauin, 1983). The outcome of this analysis is a representative measure of the conceptual structure and development of the knowledge related to an entire domain. Table 5 lists the top-10 research keywords and their occurrences. Further, the graphical visualization of research keywords co-occurrences using a networked graph reveals the relationship among the most popular research keywords. Fig. 4 visualizes the most important research keywords related to the IM domain. Apart from 'information management', the other most common keywords are 'innovation', 'sustainable development', 'economics', 'regional planning', 'industrial engineering', 'competition', and 'information systems'.

3.5. Topic modeling for thematic analysis

Following the process discussed in the methodology section, this study has analyzed the text content of each research document, including journal articles, conference papers, reviews and book chapters using STM based topic modeling. The text corpus for the topic modeling

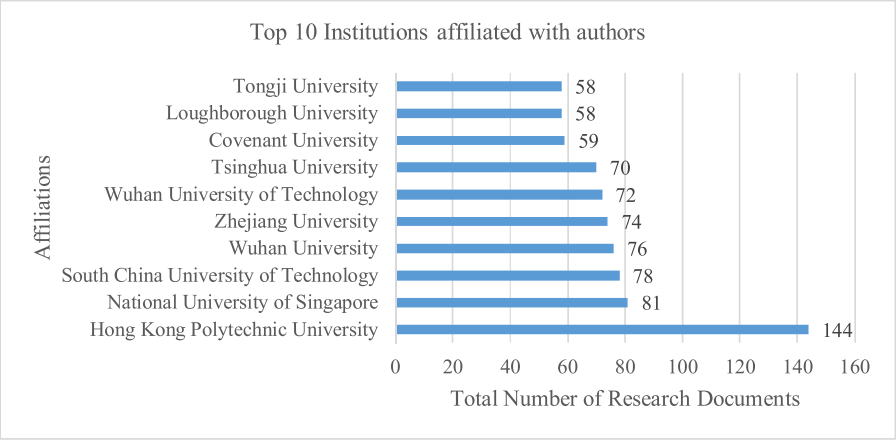


Fig. 3. Top 10 institutions affiliated with authors.

Table 4
Top 10 countries with the highest number of publications.

Country	Total Publications	Total Citations
China	4404	9786
United States	2097	97,247
Romania	1643	484
Russian Federation	1213	1978
Czech Republic	1156	766
United Kingdom	867	19,359
Malaysia	780	1212
Australia	553	8826
Germany	537	4726
Indonesia	500	353

Table 5
Top research keywords and their occurrences.

Keyword	Occurrences
Information Management	17,309
Innovation	4128
Sustainable Development	3290
Economics	2968
Regional Planning	2057
Industrial Engineering	2054
Competition	1697
Information Systems	1600
Knowledge Management	1336
Economic and Social Effects	1271

was prepared by combing document title, keyword terms, and the research abstract. This study has first analyzed all the research documents together to extract key themes related to the entire IM domain. A total of 19,916 documents were analyzed to discover the latent topics and their prevalence from 1970 to 2019. Further, this study has also done a separate document category wise analysis on research articles published in journals and research papers of conference proceedings subsequently. The most significant issue in topic modeling is to identify the correct number of latent topics. This study has used the trade-off between semantic coherence and exclusivity which is consistent with previous work related to analyzing the distinctness of extracted latent topics (Kuhn, 2018).

3.5.1. Results: all research documents

In the first phase of analysis, all 19,916 research documents related to the IM domain published from 1970 to 2019 are used to build the STM model which is configured to execute a maximum of 150 expectation-maximization iterations. The change in the approximate bound between two successive expectation-maximization iterations was carefully observed and it was found that the STM model on 19,916 documents was converged after 97 iterations of the expectation-maximization algorithm.

This study has discovered total 16 topics from a corpus of 19,916 research documents related to the IM domain using STM based on the maximum averaged held-out likelihood. The semantically descriptive topics were given meaningful labels based on highly probable words. The extracted topics are shown in Fig. 5, and topics with their expected topic proportions are shown in Fig. 6. As discussed in the methodology, this study has concatenated the most frequent bigrams to include these meaningful bigrams in the topic modeling process.

Semantic coherence and exclusivity are the two most important constructs that measure the overall quality of topic models. The average semantic coherence is a measure of co-occurrence of the most probable words in each extracted topic across all documents. In addition, exclusivity for a given topic ensures that the high-probability words for the topic have low probabilities conditional on other topics, which means the most probable words of a given topic should not appear in other topics. The average semantic coherence and exclusivity scores for all topics are given in Table 6, which shows that exclusivity is in the range of 11.33–11.80 for all the topics, and the range of semantic coherence is -191.29 to -102.45. The difference in semantic coherence values related to two topics indicates that the top words of a given topic do not co-occur equally with top words of other topics in the given corpus of documents. For example, the exclusivity scores for both Topic-1 (International Accounting and Global Business) and Topic-12 (Financial Performance and Investment) is the same (11.66) as well as for Topic-3 (Information, Web, and User) and Topic-8 (Industry and Industrial Innovation) is also identical (11.63). But the variation in the semantic coherence indicates that the top words in both topics do not occur equally within the documents. To confirm the variation in the sematic coherence for topic pairs having similar exclusivity scores, this study has investigated the topical contrasts for these topic pairs.

STM links specific words in a document corpus to latent topics based on their semantic associations, in which case, each topic can be summarized by a small collection of words. These top-*n* words can also be used to provide a label or intuitive meaning to a particular topic. For example, semantically connected words like 'International', 'Account', 'Global', 'Business', 'Trade', 'Market', 'Standard', 'Association', 'Report', and 'Information Management' form the Topic-1 where the top-3 or top-4 words that have the highest probability of occurring in the topic can be used to label the topic in a meaningful way and provide an intuitive topic definition explicitly. The method used to label the topics is purely data driven and is consistent with previous studies that have used STM on open-ended survey responses (Roberts et al., 2014), aviation incident reports (Kuhn, 2018), hotel reviews (Hu, Zhang, Gao, &

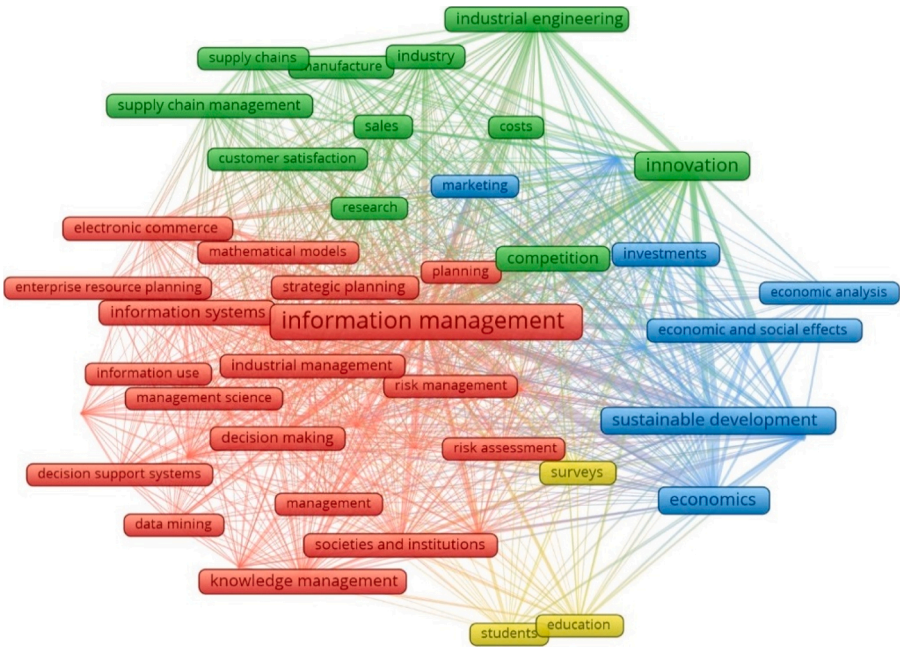


Fig. 4. Co-occurrence of author-specified keywords.

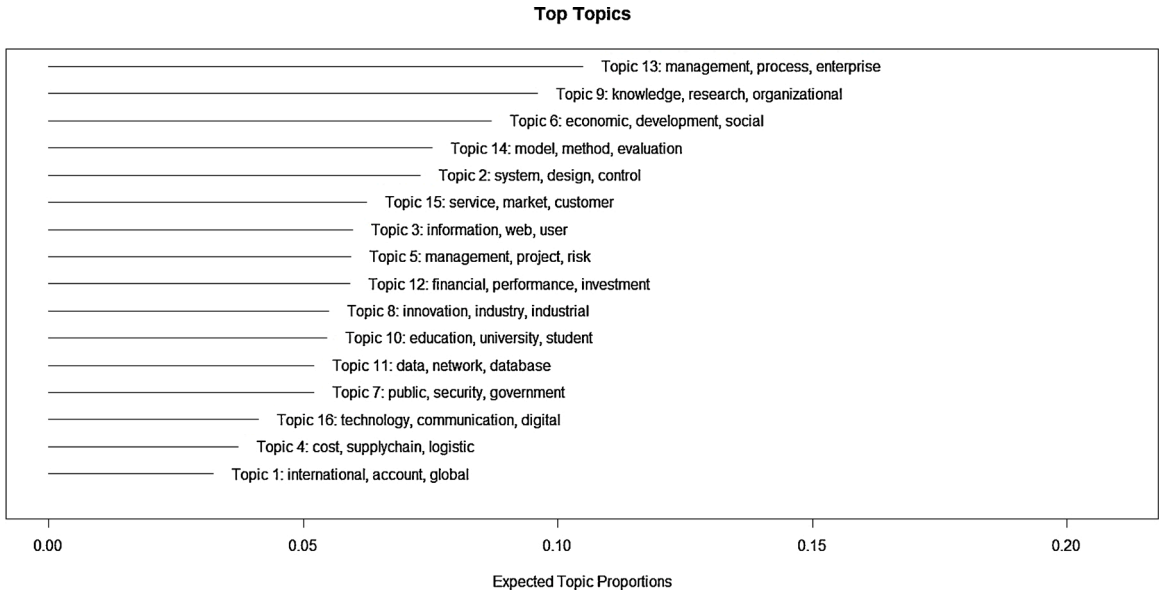


Fig. 5. Extracted topic labels from all research documents.

Bose, 2019), airlines reviews (Stamolampros, Korfiatis, Chalvatzis, & Buhalis, 2019; Stamolampros, Korfiatis, Kourouthanassis, & Symitsi, 2019), restaurants reviews (Park, Chae, & Kwon, 2018), hospitality literature (Park, Chae, & Kwon, 2018), and employees' online reviews (Stamolampros, Korfiatis, Chalvatzis et al., 2019; Stamolampros, Korfiatis, Kourouthanassis et al., 2019). Table 7 summarizes the top 10 words associated with each latent topic. Associating a topic with the words having the highest conditional probability of occurring in the topic has a significant drawback that few words such as Information, Data, System, and Information Management will present in many topics as highly probable words. The presence of such words adds little value to the topic definition and its meaningfulness. However, there are other metrics like Lift, and FREX suggested in the statistical text analysis literature which are used by many recent studies to provide more meaningful description of the latent topics (Kuhn, 2018). Lift represents only those words which are more frequent and common within a particular topic than they are across the entire document collection. In this way, Lift usually highlights rare words within a particular topic, which are not frequent across the document corpus. Another useful measure is frequency–exclusivity or FREX which highlights group of words which are both frequent and exclusive for a particular topic. The FREX ratio is used as a measure to avoid the most frequent words to describe a topic. The ratio of term frequency conditional on a topic (*tf*) to the term-topic exclusivity (*tte*) or term frequency-inverse term-topic exclusivity represents FREX. Table 7 also summarizes the top 10 words associated to each latent topic as per FREX and Lift score.

Further, the Topic-2 in Table 7 is labeled as 'System Design and Management Control', and it is formed by a probabilistic distribution over words like 'System', 'Design', 'Control', 'Management', 'Model', 'Support', 'Software', 'Information System', 'Decision', 'Computer', and

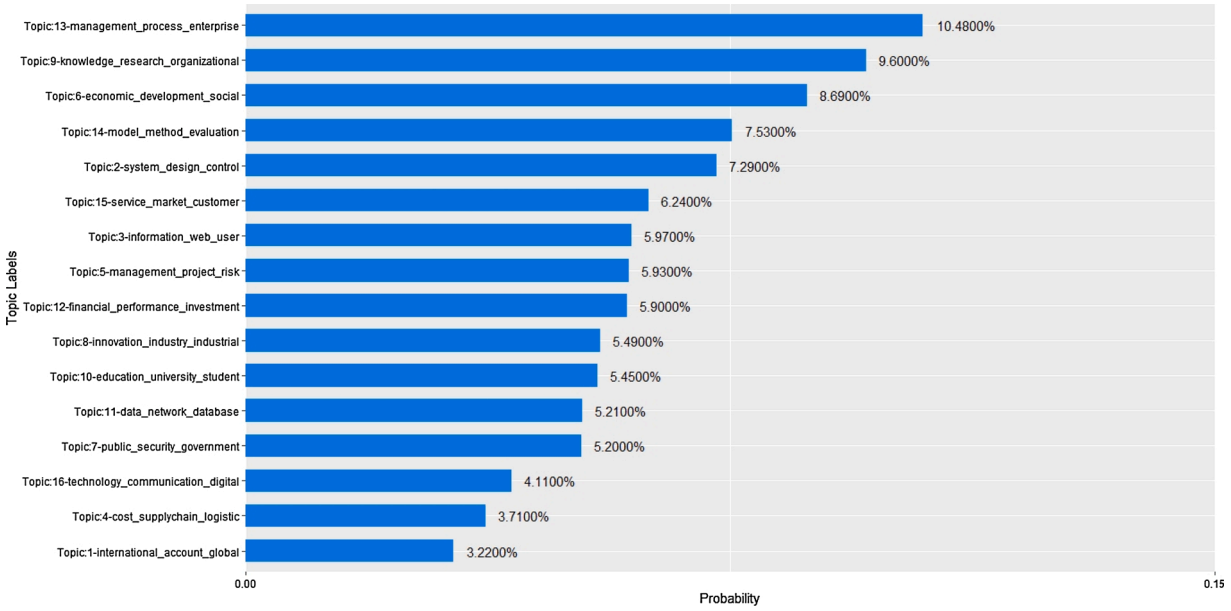


Fig. 6. Extracted topics and expected proportions (frequent bigrams are concatenated).

Table 6
Average semantic coherence and exclusivity of topics.

SN	Topic Label	Semantic Coherence	Exclusivity
1	International Accounting and Global Business	−114.79	11.66
2	System Design and Management Control	−107.88	11.67
3	Information, Web, and User	−157.50	11.63
4	Supply Chain and Logistics Cost	−191.29	11.80
5	Project Risk Management	−109.52	11.53
6	Economic and Social Development	−102.45	11.50
7	Public, Security, Government	−188.56	11.44
8	Industry and Industrial Innovation	−158.43	11.63
9	Organizational Knowledge and Research	−120.50	11.47
10	Education, Universities and Students	−139.52	11.48
11	Data, Database and Networks	−182.21	11.60
12	Financial Performance and Investment	−149.43	11.66
13	Process Enterprise Management	−117.15	11.72
14	Models, Methods, and Evaluation	−142.76	11.33
15	Service, Market, and Customer	−154.77	11.52
16	Digital Communication Technology	−179.89	11.74

many other. Similarly, the Topic-3 labeled as 'Information, Web, and user', is a probabilistic distribution over words which are semantically analogous to words like 'Information', 'Web', 'User', 'Information Management', 'Document', 'Language', 'Content', 'Search', 'Semantic', and 'Image' etc. As per the basic assumption of probabilistic generative models, each document in a text corpus is a "*distribution over a fixed vocabulary*" (Blei et al., 2003), so all the documents in the corpus can be created from a probabilistic mixture of these topics and their topical words.

The perspective visualization in Fig. 7 depicts the topical contrast for Topic-1 (International Accounting and Global Business) and Topic-12 (Financial Performance and Investment) as well as Topic-3 (Information, Web, and User) and Topic-8 (Industry and Industrial Innovation). In this topical contrast visualization, the topical words are visualized with size proportional to their probability within the topic. The topical words which are oriented more along the X-axis give an indication about how much they are semantically associated with one topic against the other.

Further, the wording difference in these topics is presented in Fig. 8. The research documents related to information management in the Topic-1 (International Accounting and Global Business) mention more words related to macro-level aspects like 'International', 'Account', 'Global', 'Business', 'Trade', 'Market', 'Standard', 'Association', 'Report', and 'Information Management'. Contrary to this, the words mentioned in the Topic-12 (Financial Performance and Investment) seem to focus on the discussion of the micro-level aspects like 'Financial', 'Performance', 'Investment', 'Company', 'Capital', 'Bank', 'Analysis', 'Corporate', 'Market', and 'Study'. Similarly, the topical contrast analysis on Topic-3 (Information, Web, and User) and topic-8 (Industry and Industrial Innovation) show that Topic-3 (Information, Web, and User) focuses more on research related to 'Web', 'User', 'Content', 'Search' and 'Semantic'. In contrast, Topic-8 (Industry and Industrial Innovation) is formed by highly frequent terms like 'Innovation', 'Industry', 'Industrial', 'Environmental', 'Development', 'Energy', 'Analysis', and 'Efficiency'.

To further investigate and quantify the association among all the extracted topics, this study has conducted a correlation analysis using an estimated marginal topic proportion correlation matrix, which reveals correlations between topics. Fig. 9 depicts a visualization of the correlations among the topics presented in Table 7. Literature suggests that a correlation value of more than 0.7 is considered a strong correlation, while the correlation value between 0.3 and 0.7 is deemed as moderate correlation (Ratner, 2009). Moreover, a positive correlation between the two latent topics shows that there are many documents in the corpus, which contain both the topics equally. The visualization of the correlations among the topics confirms that all the correlation values are less than 0.3, which indicates a weak or no correlation among the extracted topics.

To confirm the quality of the extracted topics, this study also analyzed the maximum-a-posteriori (MAP) estimates for document-topic loadings. Fig. 10 shows a histogram plot of the MAP estimates of the research article-topic loadings across all research documents of the IM domain. This graph shows the estimates of document-topic proportions, which is actually the expected distribution of topic proportions across all the research documents from 1970 to 2019. The median is denoted by a dashed red line in Fig. 10. The statistical mixture hypothesis posits that each document is a probabilistic mixture of a limited number of topics where the probability of key topics will be high and non-key topics will be low. Fig. 10 gives a rough estimate of the spread of topics across documents. This plot is very significant in STM based topic modeling because it depicts which topics are being extracted from which

Table 7
Extracted topics from STM and intuitive meanings.

No.	Topic Label	Proportion	Top-10 Words	FREX	Lift
1	International Accounting and Global Business	0.032	International, Account, Global, Business, Trade, Market, Standard, Association, Report, Information Management	Trade, Export, Currency, Internationalization, IFRS, Accountant, Account, Lease, Wine, Globalization	Accounting, Anglo, Arneburg, Beekeeper, Beekeeping, BRI, Carte, Cieszyn, Citation, CRAS
2	System Design and Management Control	0.073	System, Design, Control, Management, Model, Support, Software, Information System, Decision, Computer	SOA, DSS, Multi-Agent, Workflow, System, Agent, Automation, Subsystem, Maintenance, Cad	SOA, Anlagen, Anti-Interference, Append, Architecturally, ASPNET, AST, ATL, ATS, AUML
3	Information, Web, and user	0.06	Information, Web, User, Information Management, Document, Language, Content, Search, Semantic, Image	Ontology, Text, Semantic, Retrieval, Recommender, Watermark, Arabic, Representation, Annotation, Language	Spam, Watermark, Abet, AGRAFO, Ale, Alexandria, Alphabet, Anti-plagiarism, ARABE, ARCMAP
4	Supply Chain and Logistics Cost	0.037	Cost, Supply Chain, Logistic, Tourism, Demand, Price, Chain, Supplier, Share, Transportation	Supply Chain, Logistic, Tourism, Supply Chains, Inventory, Transport, Tourist, Destination, Reverse, Port	Abate, Agro-Touristic, AINSI, Alpine, Ancestral, Anti-Bullwhip, Antitrust, Latest, Backorder, Tourism
5	Project Risk Management	0.059	Management, Project, Risk, Construction, Information System, Assessment, Analysis, Research, Factor, Development	Project, BIM, Contractor, Risk, Construction, Maturity, Assessment, Outsource, Failure, MIS	AHF, Alabama, Anticipatory, Apo, Architecting, Benefit-cost, Cavern, Change Management, CICS, Constructability
6	Economic and Social Development	0.087	Economic, Development, Social, Economy, Regional, Country, Growth, Information Management, Analysis, Region	Rural, Agriculture, GDP, Agricultural, Tax, Farmer, Romania, Republic, Czech, Romanian	Mortgage, AAI, Aboriginal, Acreage, Actuarial, Administrative territorial, Aggravation, Agricultural, Agriculture, Agro climatic
7	Public, Security, Government	0.052	Public, Security, Government, Health, Policy, Healthcare, Management, Record, Information Management, Service	Health, Healthcare, Medical, Information Security, Hospital, E-Government, Disaster, Patient, Compliance, Disease	ABAC, AHA, ALAM, Alzheimer, Alzheimer's, Anesthesiology, Anna, Anti-infective, Antimicrobia, Antivirus
8	Industry and Industrial Innovation	0.055	Innovation, Industry, Industrial, Environmental, Development, Energy, Analysis, Efficiency, Innovative, Engineer	Green, Coal, Emission, Carbon, Ecological, Renewable, Conservation, Environmental, Waste, Circular	Bauxite, Biodiesel, Biofuel, CCR, Coal, Coalbed, Cradle, Cyber society, Diagnostic, Eco-innovation
9	Organizational Knowledge and Research	0.096	Knowledge, Research, Organizational, Study, Organization, Knowledge Management, Literature, Performance, Firm, Theory	Knowledge Sharing, Tacit, Knowledge Management, Organizational, Organizational, Leadership, Culture, Knowledge, Transformational, Inter-Organizational	Academic industry, Autopoiesis, BISL, Clevel, Co-create, Co-evolve, Coworker, Creativities, Cross boundary, Culture-commerce
10	Education, Universities and Students	0.055	Education, University, Student, Learn, Human, Employee, Study, High, Research, Work	Education, Student, Teach, Educational, School, E-learning, College, Woman, Graduate, Teacher	Achiever, Adams, Addie, Adolescent, Learning, Affectivity, Agreeableness, Al-Farabi, Anaerobic, Apprentice
11	Data, Database and Networks	0.052	Data, Network, Database, Data mining, Storage, Big Data, Management, Cloud, Propose, Distribute	Cloud, Data Warehouse, Association Rules, Encryption, Data Warehouses, SAAS, QOS, SQL, Scalability, Geospatial	ACQS Acyclic, Airborne, AODY, Application-level, Asynchrony, Attacker, Audio-video, Automatically, Availability
12	Financial Performance and Investment	0.059	Financial, Performance, Investment, Company, Capital, Bank, Analysis, Corporate, Market, Study	CSR, Loan, Intangible, Nonfinancial, Shareholder, Liquidity, Estate, Bank, Volatility, Dividend	Dow, Microfinance, Acquirer, Altman, Altmans, Ratio, Amortization, ANCOVA, Anti-takeover, Application
13	Process Enterprise Management	0.105	Management, Process, Enterprise, Business, Product, Resource, Quality, Strategy, Company, Plan	ERP, Business Process, Lean, Reengineer, Business Processes, Strategic, Competitive, Manufacture, Enterprise, TQM	MRP, APS, Asset-Intensive, BAE, Baldridge, Baldrige, Bandage, Bathroom, BDIM, computer-aided information systems
14	Models, Methods, and Evaluation	0.075	Model, Method, Evaluation, Algorithm, Analysis, Fuzzy, Optimization, Problem, Time, Result	Fuzzy, Neural, Genetic, Entropy, Stochastic, Heuristic, Rough, Nonlinear, Bayesian, Noise	ACO, Anneal, ANNs, ANSYS, Ant-colony, AOA, BCD, Bilinear, Bility, Brownian
15	Service, Market, and Customer	0.062	Service, Market, Customer, Online, Study, Consumer, Factor, Research, Behavior, Model	Brand, E-Commerce, Social Media, Loyalty, Advertise, Service Quality, CRM, Buy, Technology Acceptance, Auction	Auction, Chat, Adoptability, Advert, Agadir, Arousal, Attitude-Behavior, August-December, Benevolence, Bitner
16	Digital Communication Technology	0.041	Technology, Communication, Digital, Mobile, Internet, Application, Virtual, Electronic, ICT, Information Management	RFID, Radio, Block-Chain, Cellular, Thing, Mobile, Smart City, ICT, Smart, Virtual	Industry, Non-contact, Radio-frequency, Video-conferencing, Barcode, RFID, Aero-Spatial, AIDC, Anti-counterfeiting, Anytime-Anywhere Computing

documents. It is evident from the plot that each extracted latent topic has no relation or very little relation with several research documents.

Fig. 11 visualizes the covariate effect estimation of the topic prevalences over the years starting from 1970 to 2019. The dotted lines represent the confidence interval in the plot. It can be inferred that the topic prevalences gradually change year to year. As evident from the visualization that Topic 1- 'International Accounting and Global Business' and Topic 10-'Education, Universities, and Students' are showing a stable or rather declining trend till 2010 and after this year, the trend is rising. The same trend is visible for Topic 6-'Economic and Social Development' with 2006 as the trend change point. However, the Topic 2-'System Design and Management Control' is showing a rising trend until the year 1990, after which the trend is declining. The Topic 9-'Organizational Knowledge and Research' is also showing a rising trend during 1970–2000, after which the trend is declining. The same trend is visible for Topic 16-'Digital Communication Technology' where a change point is visible around the year 2005.

The Topic 3-'Information, Web, and User' and Topic 5-'Project Risk Management' are showing a gradual decline over the entire period 1970 to 2019. The Topic 4-'Supply Chain and Logistics Cost' and Topic 8-'Industry and Industrial Innovation' are showing a contrasting trend compared to other topics where a rise in the research trend is evident till 2010, and then a declining trend is recorded till 2019. Further, Topic 7-'Public, Security, Government' and Topic 11-'Data, Database and Networks' have a stable trend of research.

Another trend worth reporting is that Topic 12-'Financial

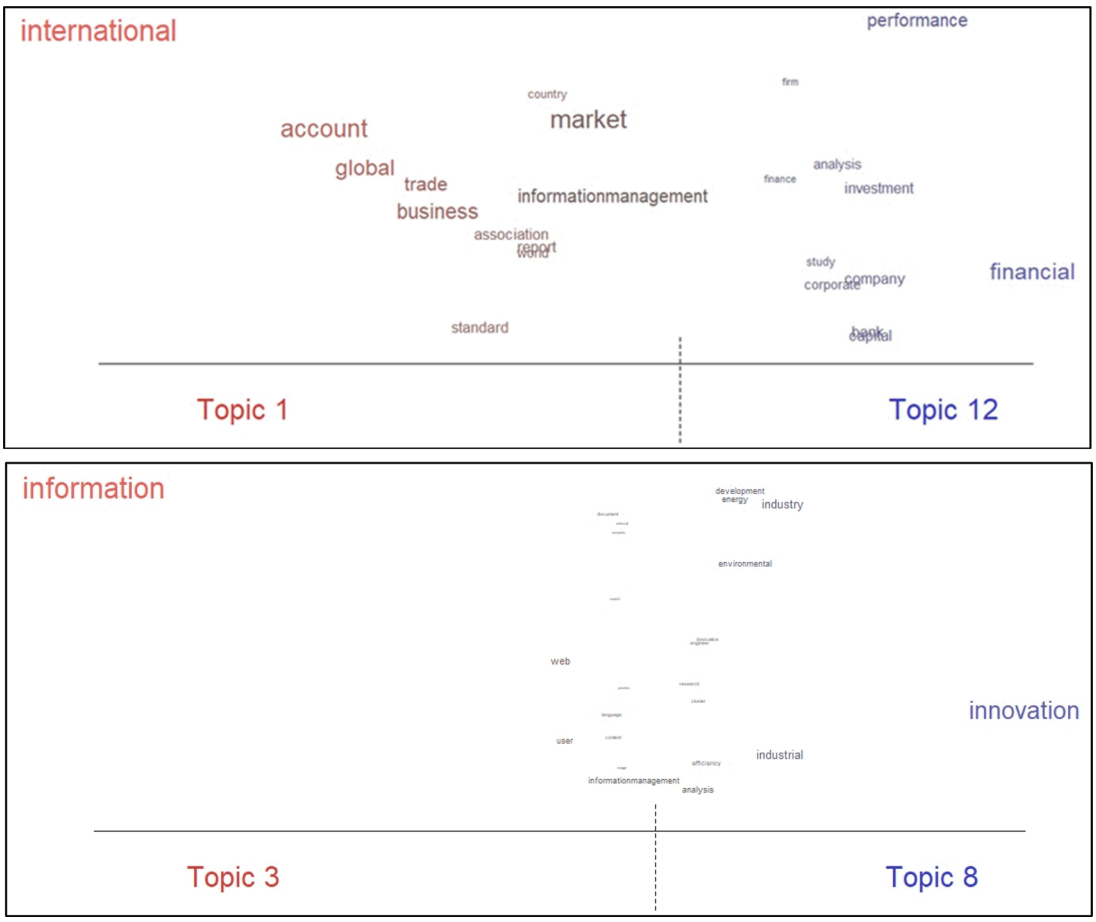


Fig. 7. Topical contrasts in two topic pairs having similar exclusivity scores.

Performance and Investment' is formed by words like 'Financial', 'Performance', 'Investment', 'Company', 'Capital', 'Bank', 'Analysis', 'Corporate', 'Market', 'Study' etc. and the trend indicates that this theme was almost stable during 1970–2010 due to which there is no visible pattern, and the trend is almost flat. However, this topic has witnessed a change point around the year 2010, and after this, the trend is increasing.

Further, the trend related to Topic 14-'Models, Methods, and Evaluation' is different where a rising trend can be observed over the years 1970–2010, and after the change points around 2010, a decreasing trend can be observed over the years. The increasing trend reveals the growing academic interest of contributors to the IM domain for these themes. In contrast, the decreasing trend unveils that researchers are attracting towards other themes which are emerging such as the Web, social media, international business, finance, and global economic related research.

3.5.2. *Results: conference papers*

Academic and business research conferences are a great medium of knowledge dissemination. As IM is a highly dynamic domain, top tier conferences in this domain are highly prestigious among researchers striving to present their most relevant research accomplishments. To perform a separate analysis on conference papers related to IM domain, all 14,560 research documents published in various conferences from 1970 to 2019 are used to build the structural topic model.

This study has discovered total 16 topics from a corpus of 14,560 research documents related to the IM domain using STM based on the maximum averaged held-out likelihood. The semantically descriptive topics were given meaningful labels based on highly probable words. The extracted topics are shown in Fig. 12, and topics with their expected topic proportions are shown in Fig. 13. As discussed in the methodology, this study has concatenated the most frequent bigrams to include meaningful, frequent bigrams in the topic modeling process.

The average semantic coherence and exclusivity scores for all topics are given in Table 8. As shown in Table 8, exclusivity is varying from 11.18 to 11.80 for all the topics, and the range of semantic coherence is –217.62 to -98.89. As discussed earlier, the difference in semantic coherence values related to two topics indicates that the top words of a given topic do not co-occur equally with top words of other topics in the given corpus of documents.

Similar to the analysis in the previous Section for quantifying the association among all the extracted topics from conference papers, this study has conducted a correlation analysis using an estimated marginal topic proportion correlation matrix, which reveals correlations between topics extracted from conference papers published from 1970 to 2019. Fig. 14 depicts a visualization of the correlations among the topics presented in Table 9. The correlations plot confirms that all the correlation values are less than 0.3, which indicates a weak or no correlation among the extracted topics. Further, to confirm the quality of the extracted topics, this study also analyzed the maximum-a-posteriori (MAP) estimates for document-topic loadings for all conference papers. Fig. 15 shows a histogram plot of the MAP estimates of the research article-topic loadings across all research documents of the IM domain. Fig. 15 gives a rough estimate of the spread of topics across documents. It is evident from the plot that each extracted latent topic has no relation or very little relation with several conference papers.

Fig. 16 visualizes the covariate effect estimation of the topic prevalences over the years starting from 1970 to 2019 for all conference papers. The dotted lines represent the confidence interval in the plot. It can be inferred that the topic prevalences gradually change year to year. As

Fig. 8. Word cloud visualizations of topic pairs: Topic 1 and 12, and Topic 3 and 8.

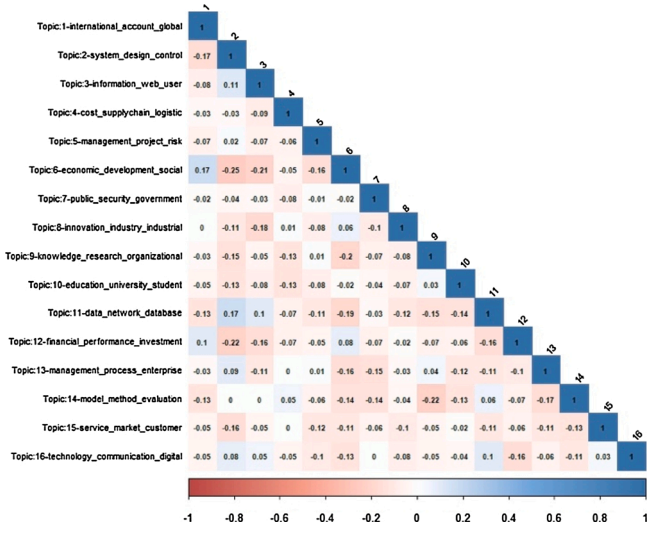


Fig. 9. Correlations among topics.

evident from the visualization that Topic 1-'Economic, Development, and Countries', Topic 3-'Social Public Study', Topic 5-'Market and Customer, Service', Topic 8-'Banks and Financial Investments', and Topic 12-'Energy and Environment' are showing a growing trend during the period. However, an opposite trend is visible for Topic 2-'Information Management and System Management', Topic 10-'Digital Data and Service', Topic 13-'Model, Process, Quality', and Topic 15-'Education, Student and Learning'. However, the rest of the themes are stable over the period.

3.5.3. Results: journals' articles

A peer-reviewed journal publication is generally considered superior, especially with a good impact factor and journal articles are usually more elaborative than the conference papers. The research content, formal structure of knowledge presentation, and the quality of the rigorous peer-review process justify a separate analysis of research articles of IM domain published in various journals. Hence, this study has performed a separate analysis of all 4735 research articles published in various journals from 1970 to 2019. The content of these research articles is used to build a separate topic model using STM. This study also supports the observation that some top conferences in the IM domain also publish their complete proceedings as journal issues, in which case there are chances that a few research documents may be repeated in both conferences and journals. However, data-deduplication at this mass level was not feasible and out of the scope of this work.

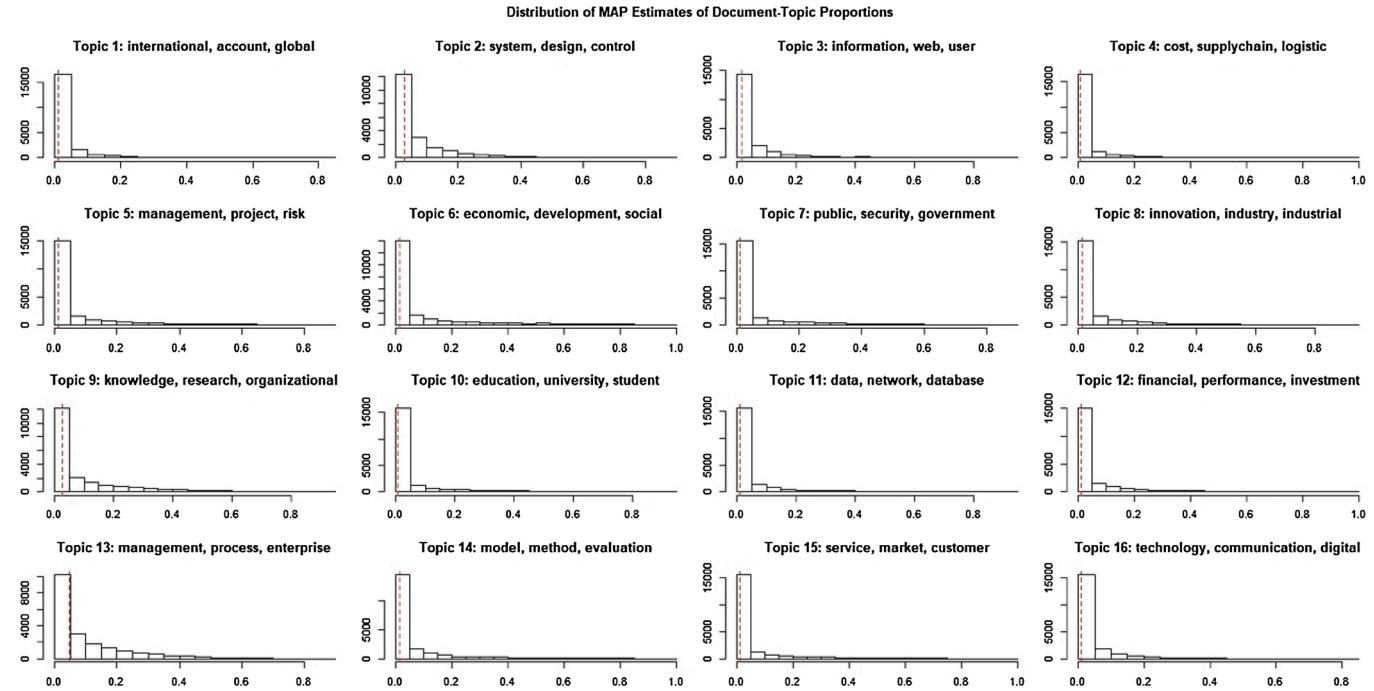


Fig. 10. Topic proportions within documents.

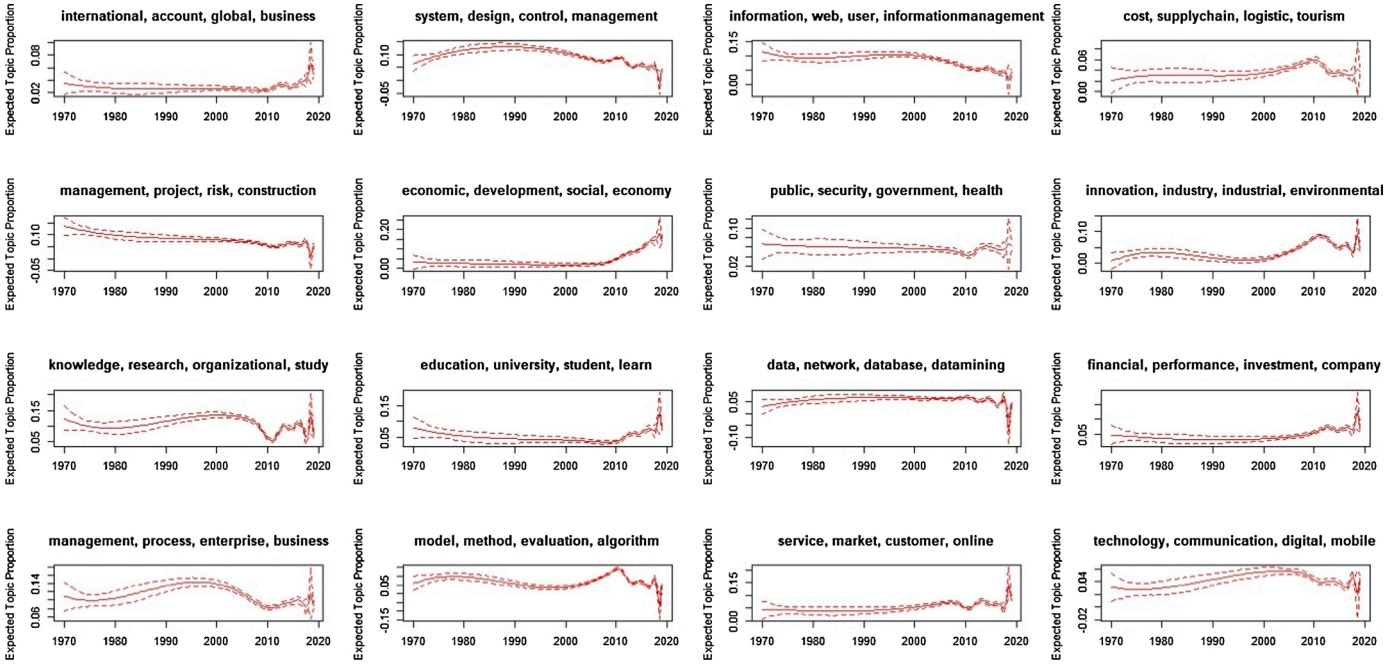


Fig. 11. Covariate effect estimation of the topic prevalences over publication years.

To provide a fair comparison with previous analyses on all research documents and conference papers, this study has discovered total 16 topics from a corpus of 4735 research documents related to the IM domain using STM based on the maximum averaged held-out likelihood. The semantically descriptive topics were given meaningful labels based on highly probable words. The extracted topics are shown in Fig. 17, and topics with their expected topic proportions are shown in Fig. 18. Similar to the previous Sections, this study has concatenated the most frequent bigrams to include meaningful, frequent bigrams in the topic modeling process.

The average semantic coherence and exclusivity scores for all topics are given in Table 10, which shows that exclusivity is varying in the range of 11.08–11.83 for all the topics, and the range of semantic coherence is -180.59 to -85.94.

Similar to the analysis in the previous Section for quantifying the association among all the extracted topics from the journal articles, this study has conducted a correlation analysis using an estimated marginal topic proportion correlation matrix, which reveals correlations between topics extracted from journal articles related to IM domain published from 1970 to 2019. Fig. 19 depicts a visualization of the correlations among the topics presented in Table 11. Fig. 19 confirms that all the correlation values are less than 0.3, which is an indicator of a weak or no

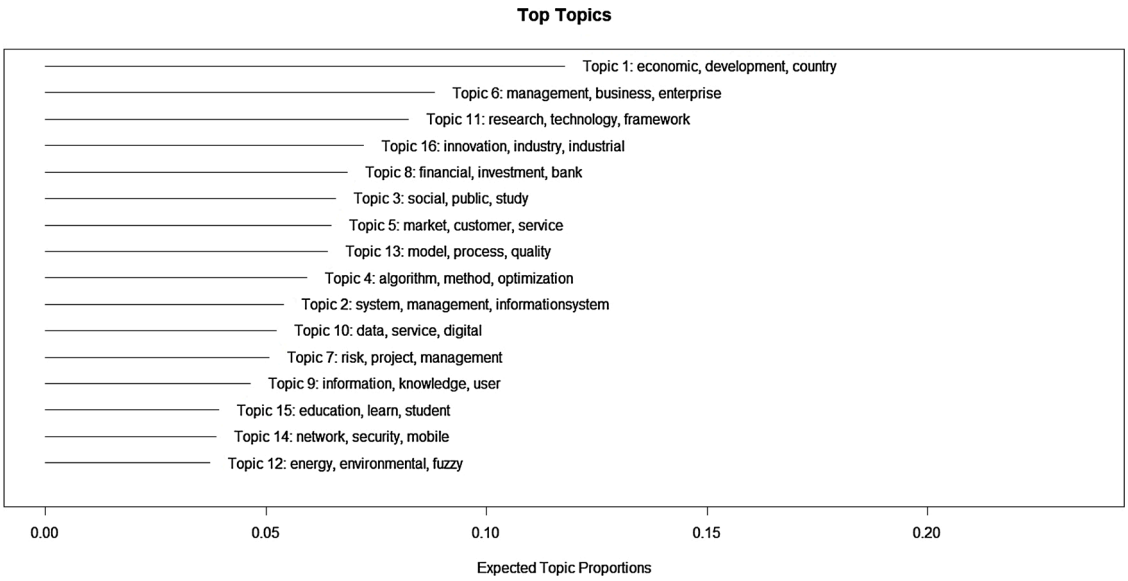


Fig. 12. Extracted topic labels from conference papers.

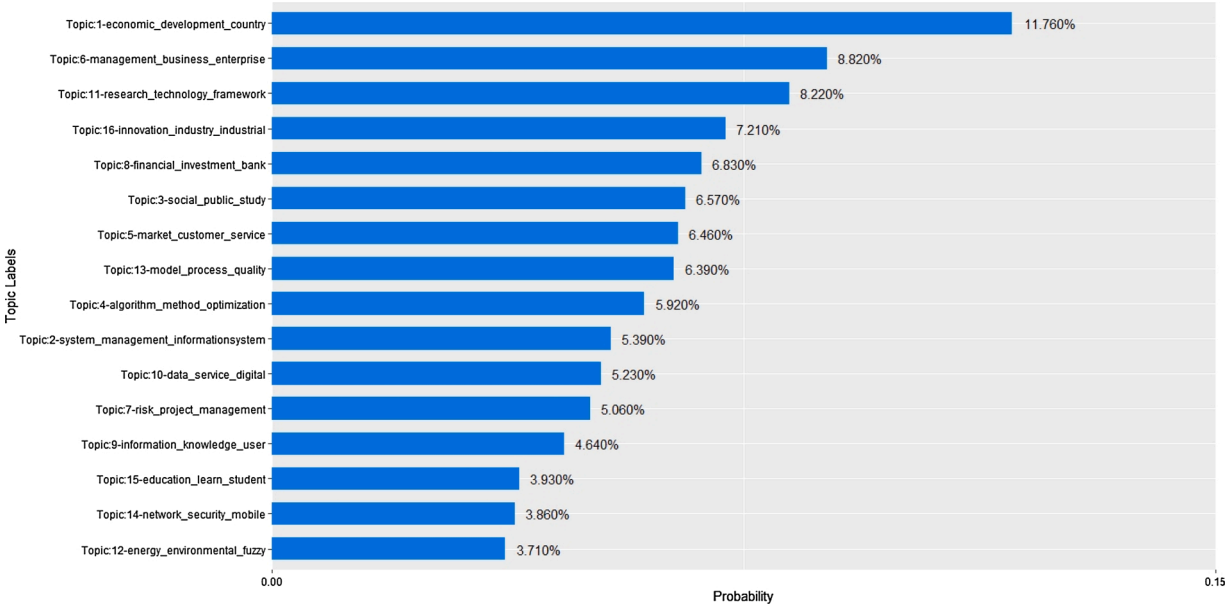


Fig. 13. Extracted topics and expected proportions (only conference papers).

correlation among the extracted topics.

Further, to confirm the quality of the extracted topics, this study analyzed the maximum-a-posteriori (MAP) estimates for document-topic loadings for all conference papers. Fig. 20 shows a histogram plot of the MAP estimates of the research article-topic loadings across all journal articles of the IM domain. Fig. 20 gives a rough estimate of the spread of topics across documents. It is evident from the plot that each extracted latent topic has no relation or very little relation with several journal articles.

Fig. 21 visualizes the covariate effect estimation of the topic prevalences over the years starting from 1970 to 2019 for all journal articles. The dotted lines represent the confidence interval in the plot. It can be inferred that the topic prevalences gradually change year to year. As evident from the visualization that Topic 2- 'Environmental Management and System', Topic 5-'Public Record Information', Topic 6-'Performance Quality Management', Topic 7-'Social Communication and Trust', and Topic 9-'Knowledge, Innovation, and Knowledge management' are showing a growing trend during the period. However, a contrasting trend is visible for Topic 1-'Logistics and International Growth', Topic 3-'Team, Role, and Manager', Topic 4-'Information Management and Organization', Topic 11-'Models and Systems for Decisions', and Topic 12-'Information, Cost, Market'.

Another trend worth reporting is that the Topic 10-'Information System and Database' and Topic 13-'Product and Enterprise Management' have a growing trend till 1990 and 2002, respectively, and after these change points these research themes have witnessed a decline in research trends. Another trend is related to Topic 16- 'Research, Management and Risks' for which there is a declining trend till the year 2000, but after that, this theme is observing growth in the total number of publications.

4. Discussion

This study reports the evolution of scientific research related to the

Table 8
Average semantic coherence and exclusivity of topics from conference papers.

No.	Topic Label	Semantic Coherence	Exclusivity
1	Economic, Development and Countries	−98.890	11.34
2	Information Management and System Management	−145.17	11.67
3	Social Public Study	−121.24	11.18
4	Algorithm, Method, and Optimization	−155.09	11.32
5	Market and Customer, Service	−167.63	11.41
6	Management of Business and Enterprise	−121.37	11.80
7	Project Management Risks	−134.45	11.50
8	Banks and Financial Investments	−145.94	11.48
9	Information, User and Knowledge Management	−178.22	11.77
10	Digital Data and Service	−136.38	11.53
11	Research, Technology, Framework	−124.93	11.40
12	Energy and Environment	−217.62	11.79
13	Model, Process, Quality	−133.22	11.64
14	Mobile Network and Security	−189.76	11.73
15	Education, Student and Learning	−142.92	11.72
16	Industry and Industrial Innovation	−113.66	11.74

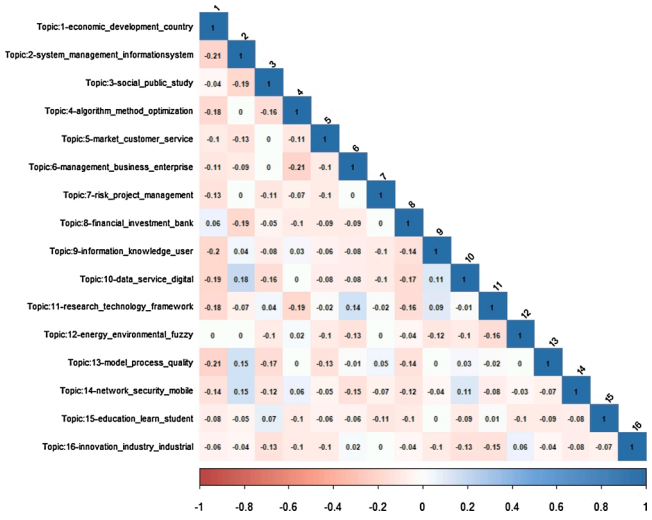


Fig. 14. Correlations among topics (only conference papers).

IM domain for the last fifty years, from 1970 to 2019. This study has analyzed a total of 19,916 research publications, including empirical research papers, case studies, book chapters, theoretical and conceptual articles, systematic literature reviews, and research meta-analyses. Moreover, conference papers and journal articles are analyzed separately, and key sixteen latent themes are identified and visualized.

Understanding, analyzing, reporting, and visualizing the evolution of scientific research related to information management published in the last fifty years was a much needed research endeavor to appreciate the extent of domain knowledge accumulated over the years. This study has analyzed the overall knowledge structure and its evolution with both retrospective and prospective analysis purposes. The publication trends, as per authors, institutions, and countries, are explored and reported. Further, this study has assessed the keywords patterns. The thematic links among the most cited and frequently co-occurring keywords are visualized using a key co-occurrence network. Finally, the underlying themes and latent topics are extracted and visualized in this study for the entire IM domain. In this way, this study summarizes and visualizes the evolution of the IM domain by identifying key publications, authors, and keywords and exploring how specific themes have evolved over time.

4.1. Discussion on the results of bibliometric analysis

Over the period, the research interest in the IM field has grown tremendously, and the last two decades have witnessed an abundance of path-breaking research works in this field. This study shows a significant increase in the number of studies and publications, specifically from 2000 to 2019. Several other research domains and subdomains have converged into IM because of which this domain has become multidisciplinary in a real sense. This is evident from the fact that the top 10 most cited articles in the IM domain are related to multidisciplinary research like knowledge management (Alavi & Leidner, 2001; Hansen et al., 1999), design science for information systems (Peffers et al., 2007), enterprise information systems (Davenport, 1998), user acceptance of information systems (Van der Heijden, 2004), role of information technology in firms (Sambamurthy, Bharadwaj, & Grover, 2003), nature of theory in information systems (Gregor, 2006), managerial decision-making tools in information systems research (Okoli & Pawlowski, 2004), information needs of top management (Rockart, 1979) and common method variance in information systems research (Malhotra, Kim, & Patil, 2006). The analysis of other top-cited articles of IM domain from 1970 to 2019 and their research keywords reveals that this field has attracted the interest of scholars and researchers from diverse areas ranging from knowledge management, enterprise systems, management information systems, enterprise resource management, business process management, electronic commerce, innovation, sustainable development, information systems security, e-government, data quality management, social commerce, industrial engineering, economic and social effects, and supply chain management etc. A comprehensive overview of the leading trends and patterns of publication highlights that the most recurrent research keywords from 1970 to 2019 are 'information management', 'innovation', 'decision making', 'sustainable development', 'economics', 'information systems', 'knowledge management', and 'economic and social effects', 'decision making, 'information technology', and 'project management'. The keyword analysis presented in Table 12 lists top-10 keywords and their frequencies from each ten years slot's research documents from each category (conference and article) for the period 1970–2019. As keywords provide vital insight into the content of the research documents, the keyword analysis provides detailed comparisons between subsamples for the periods 1970–79, 1980−89, 1990–99, 2000–09, and 2010–19. This also reveals the potential interactions between conference papers and research articles across the years.

The most prominent institutions contributing to the IM domain are Hong Kong Polytechnic University, National University of Singapore, South China University of Technology, Wuhan University, Zhejiang University, Wuhan University of Technology, Tsinghua University, Covenant University, Loughborough University, and Tongji University. These universities have executed the maximum number of research studies in this domain. The top countries as per the number of research documents published are China, the United States, Romania, Russian Federation, Czech Republic, the United Kingdom, Malaysia, Australia, and Germany.

4.2. Discussion on the results of structural topic modeling

The study also analyzes the text content of each information management research document published from 1970 to 2019 using structural topic modeling by which the semantically descriptive topics are extracted and given meaningful labels. The extraction of latent research themes identified by sophisticated text analytics based structural topic modeling enables potential researchers and practitioners to gain deeper and detailed insights from the collected literature. These themes are extracted using the advanced STM methodology which is based on modeling the hidden common textual content in research literature documents using a probabilistic generative model. In the first phase of analysis, all 19,916 research documents related to the IM domain

Table 9
Extracted topics and intuitive meanings-conference papers.

No.	Topic Label	Proportion	Top-10 Words	FREX	Lift
1	Economic, Development, and Countries	0.118	Economic, Development, Country, Economy, International, Information Management, Regional, Social, Policy, Growth	Union, Agriculture, Russia, Territory, European, Russian, Farmer, Rural, Agricultural, Europe	Cereal, TSE, Territory, Abolition, Accede, Accession, Additivity, Administrative Territorial, Aggravation, Agriculture
2	Information Management and System Management	0.054	System, Management, Information System, Control, Support, Design, Information, Decision, Monitor, Integrate	Emergency, Multi-Agent, Maintenance, Agent, Warn, System, DSS, Disaster, Traceability, Monitor	Anti-interference, Business Processing, Crowdsourced, E-Agriculture, EWS, GAIA, GDSS, granary, Haiti
3	Social Public Study	0.066	Social, Public, Study, Employee, Work, Factor, Culture, Information Management, Effect, Research	CSR, Woman, Transformational, Responsibility, Commitment, Entrepreneur, Leadership, Saudi, Family, Malaysia	Adolescence, Agreeableness, Alertness, Alzheimer, Alzheimer's, Amenity, Anybody, Autocratic, BAIA
4	Algorithm, Method, and Optimization	0.059	Algorithm, Method, Optimization, Problem, Result, Time, Neural, Propose, Image, Information Management	Algorithm, Genetic, Rough, Heuristic, Wavelet, Multi Objective, Particle, Association Rules, Watermark, Classifier	Adaboost, Agglomerative, Algorithmic, Anneal, Archaeological, ASCII, Bayes, BCD, Bilinear, Bi-Orthogonal
5	Market and Customer, Service	0.065	Market, Customer, Service, Consumer, Online, Tourism, Study, Health, Brand, Information Management	Brand, Healthcare, Social Media, Advertise, Customer Relationship, Technology Acceptance, Facebook, Halal, E-health, Consumer	Advertise, Customer Relationship, E-Banking, Halal, Repurchase, Adoptability, Advert, Advertisement, Agadir, AINSI
6	Management of Business and Enterprise	0.088	Management, Business, Enterprise, Resource, Company, Organization, Strategy, Performance, Human, Strategic	SMEs, ERP, Strategic, Competitive, Medium Sized, Organization, Enterprise, Sustainability, Business, SME	Adaptiveness, Aps, Audit, BSS, biasness, BIS, CAAS, Capability-Driven, Cloud-Enabled, Company-Corporation
7	Project Management Risks	0.051	Risk, Project, Management, Cost, Supply Chain, Construction, Assessment, Control, Information Management, Analysis	Supply Chain, Contract, Inventory, Procurement, Supply Chains, Risk, BIM, Supplier, Contractor, Game Theory	Nash, Principal-Agent, Supply-Chain, ABMS, allegro, anxious, APO, ASC, BAI, BAZSOV
8	Banks and Financial Investments	0.068	Financial, Investment, Bank, Market, Capital, Analysis, Company, Study, Finance, Regression	Loan, Cash, Earnings, Shareholder, Causality, Liquidity, Co-Integration, IFRS, Nonfinancial, Stock	CPEC, Goodwill, Microfinance, Accrual, Accrue, Acquirer, Altman, Altman's, Amman, Amortization
9	Information, User and Knowledge Management	0.046	Information, Knowledge, User, Knowledge Management, Web, Language, Share, Semantic, Content, Document	Ontology, Tacit, Semantic, Knowledge Management, Knowledge Sharing, Arabic, Library, Knowledge, Document, Language	Ontological, Agrafo, Allergy, Ancestor, Ancestral, Antiplagiarism, Arabic, Bahasa, Bioethics, Bioethical
10	Digital Data and Service	0.052	Data, Service, Digital, Architecture, Database, Application, Management, Process, Storage, Technology	cloud, big data, analytics, SOA, data warehouse, block chain, schema, cloud computing, storage, E-government	analytics, availability, big-data driven, Blockchain's, cloud sourcing, dado, data related, data store, dbas, DSMSS
11	Research, Technology, Framework	0.082	Research, Technology, Framework, Study, Literature, Review, Information System, Development, Information Management, Approach	Review, Collaboration, Literature, ICT, Organization, Journal, Organizational, Conceptual, Adoption, Community	E-Registration, Gov, ISG, Requisites, Scopus, Selection, Bibliometric, Scrum, Sociotechnica, Acclaim
12	Energy and Environment	0.037	Energy, Environmental, Fuzzy, Power, Price, Forecast, Water, Green House, Consumption	Energy, Water, Gas, Coal, Emission, Carbon, Renewable, Conservation, Electricity, Green	Aerodynamic, ANFIS, BAO, Biodiesel, Buzz, Calcaric, Chill, Chip, Coil, Comet
13	Model, Process, Quality	0.064	Model, Process, Quality, Design, Product, Software, Engineer, Production, Manufacture, Dynamic	Petri, Assembly, Diagram, Cycle, Lean, PLM, UML, Lifecycle, Software, Notation	Adoxx, Aristaflow, Artifact-Centric, Axiomatic, BAE, CATIA, CME, Co-Design, Co-Designs, CPN
14	Mobile Network and Security	0.039	Network, Security, Mobile, Internet, Communication, Transportation, Traffic, Transport, Device, Telecommunication	Mobile, Traffic, Wireless, Authentication, Congestion, Cellular, Packet, Telecom, Road, Protocol	AOA, ATN, bandwidth efficient, boat, bpsk, car-sharing, CDMA, cluster-head, CMMB, cryptanalysis
15	Education, Student and Learning	0.039	Education, Learn, Student, University, High, Train, Institution, Teach, Educational, Professional	Education, Student, University, Teach, Educational, E-Learning, School, College, Teacher, Graduate	classroom, mosque, postgraduate, accreditation, achiever, addie, alfarabi, anglo, apprentice, architectonics
16	Industry and Industrial Innovation	0.072	Innovation, Industry, Industrial, Evaluation, Analysis, Development, Logistic, Research, Engineer, Technology	Patent, DEA, High-Tech, Innovation, Logistic, Industrial, Evaluation, Province, Innovative, Industry	Agglomeration, AHM, Analysis, Analyses, Arc-view, Auto Parts, BASAL, BCC, Bertrand, Bionics

published from 1970 to 2019 are used to build the STM model and extract total 16 latent topics that are given semantically descriptive labels. To assess the quality of the topics, this study has performed topic correlation analyses and studied the trade-off between semantic coherence and exclusivity.

It is imperative to consider that topic modeling is an unsupervised learning technique, and there is no standard gold data for testing and validation, making it impracticable to validate the output of a topic model. Semantic coherence is a criterion proposed in a previous work on topic modeling (Mimno, Wallach, Talley, Leenders, & McCallum, 2011) which showed that it correlates well with the human judgment of topic quality. The FREX metric proposed by Bischof and Airoldi (2012) measures exclusivity in a way that balances word frequency (Bischof & Airoldi, 2012). Models with fewer topics have higher semantic coherence, due to which the researcher needs to look at both semantic coherence and exclusivity of words to topics. To understand these two measures and their application in highlighting the distinctness of the topics, we can plot the semantic coherence and exclusivity scores of the topics (see Table 6) as shown in Fig. 22.

The plot is based on the previous works on STM (Roberts et al., 2014, 2016), and the same has also been used in selecting the right number of topics by Kuhn (2018). On the exclusivity dimension in Fig. 22, the relative differences are small, which indicate that the top words for the topics can appear within the top words of other topics. On the semantic coherence dimension, the relative differences are larger. It indicates that the words that are most associated with the corresponding topics do not occur equally within the documents. For example, the difference in the exclusivity of Topic 3- 'Information, Web, and User' and Topic 8-

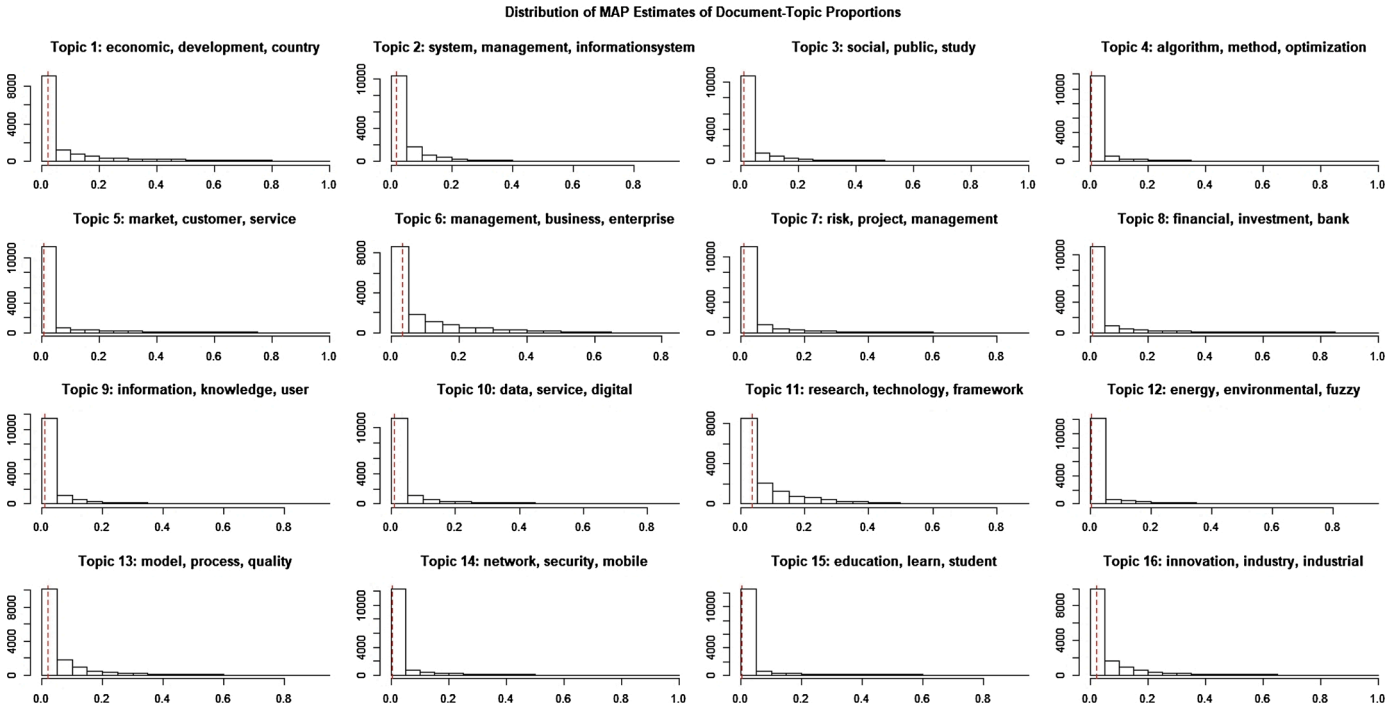


Fig. 15. Topic proportions within documents- conference papers.

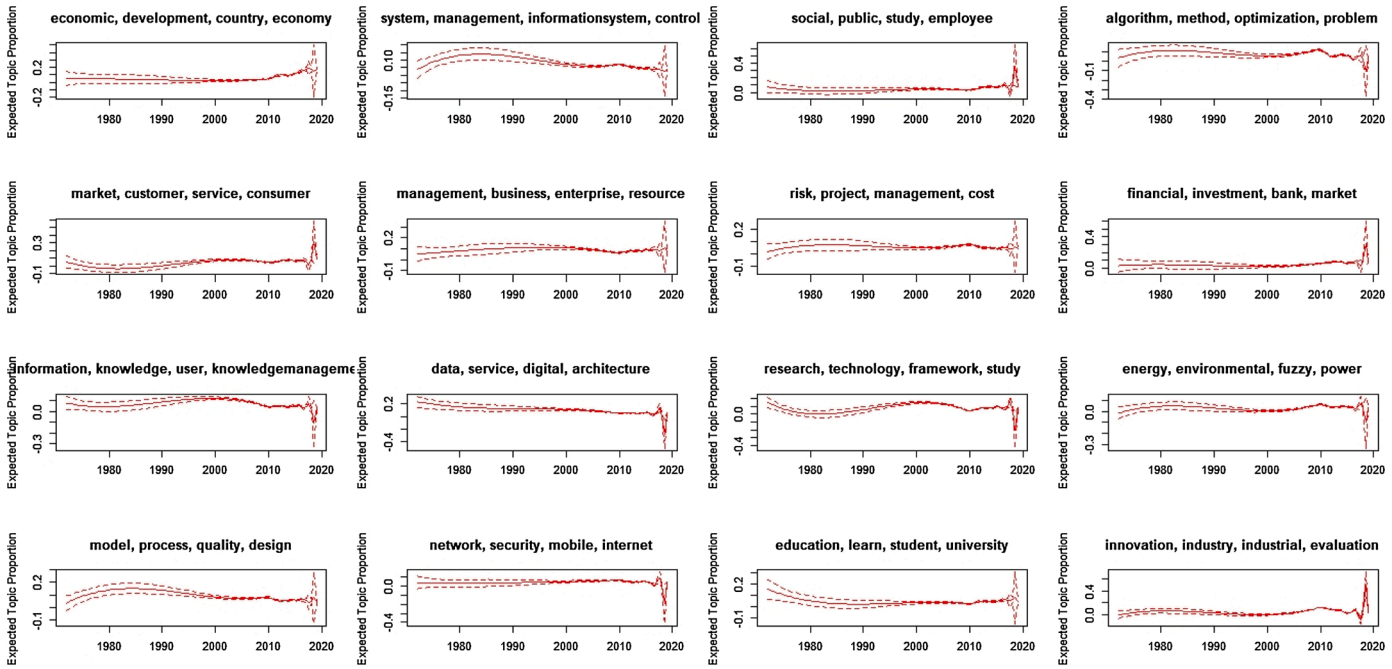


Fig. 16. Covariate effect estimation of the topic prevalences over conference years.

'Industry and Industrial Innovation' is zero. It specifies that the few words in Topic 3 may reappear in Topic 8 and vice versa. The distinctness of the topics can be measured by the distance of these data points. For example, Topic 4- 'Supply Chain and Logistics Cost' is relatively far away from Topic 14- 'Models, Methods, and Evaluation', which means that these two topics are distinct in nature.

The covariate effect estimation of the topic prevalences from all the research documents published in the IM domain over the years starting from 1970 to 2019 (Fig. 11) reveals that topics such as 'International Accounting and Global Business', 'Economic and Social Development', 'Industry and Industrial Innovation', 'Education, Universities and Students', 'Financial Performance and Investment', and 'Service, Market, and Customer' are showing a rising trend in research. However, the topics like 'System Design and Management Control', 'Information, Web, and User', 'Project Risk Management', and 'Public, Security, Government' are showing a declining trend. The useful insights related to interesting research themes and topics offered in this study have strong implications for further research in this area.

This study has also performed a separate analysis of conference papers related to the IM domain published from 1970 to 2019. A total of

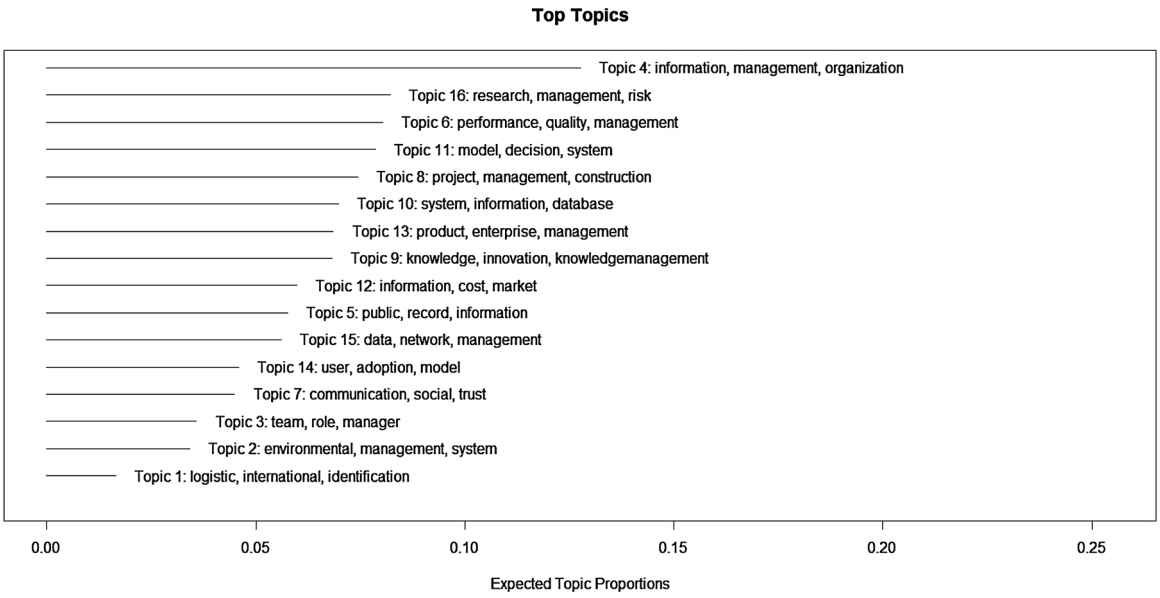


Fig. 17. Extracted topic labels from journal articles.

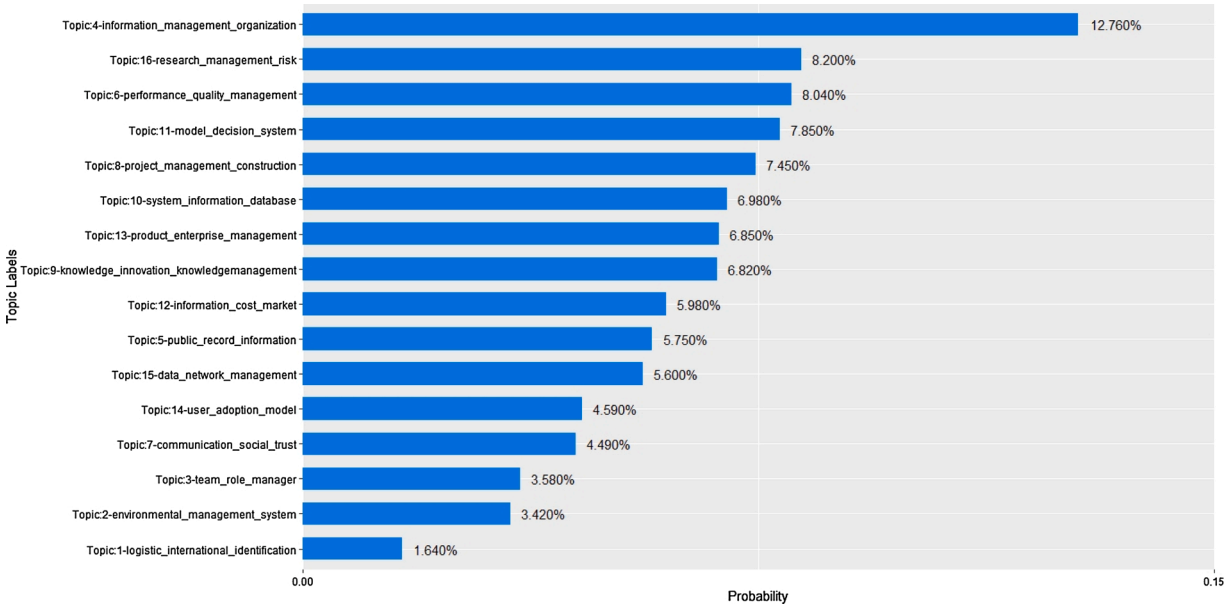


Fig. 18. Extracted topics and expected proportions (only journal articles).

14,560 research documents are published in various conferences during this period, which are used to build the STM model and extract total 16 topics from a corpus of 14,560 research documents. The extracted 16 latent topics using STM are given semantically descriptive labels, and to quantify the association among all the extracted topics, a correlation analysis is performed, and the semantic coherence and exclusivity scores of the topics are plotted (Fig. 23) to visualize the quality of topics.

The covariate effect estimation of the topic prevalences over the years starting from 1970 to 2019 for all conference papers highlights that topic 'Economic, Development, and Countries', 'Social Public Study', 'Market and Customer, Service', 'Banks and Financial Investments', and 'Energy and Environment' are showing a growing trend during the period. However, topics such as 'Information Management and System Management', 'Digital Data and Service', 'Model, Process, Quality', and 'Education, Student and Learning' are having an inverse trend.

Finally, this study has analyzed a total of 4735 research documents of IM domain published in various peer-reviewed journals indexed by Scopus. To provide an impartial evaluation consistent with the previous analyses on all research documents and conference papers, this study has extracted the same number of topics using STM from a corpus of 4735 research documents. A total of 16 semantically descriptive topics are extracted and given meaningful labels based on highly probable words. The semantic coherence and exclusivity scores of the topics are represented in Fig. 24 to visualize the quality of topics.

The covariate effect estimation of the topic prevalences over the years starting from 1970 to 2019 for all journal articles related to the IM domain reveals that 'Environmental Management and System', 'Public Record Information', 'Performance Quality Management', 'Social Communication and Trust', and 'Knowledge, Innovation, and Knowledge management' are showing a growing trend during the period. However, a contrasting trend is visible for topics such that 'Logistics and

Table 10
Average semantic coherence and exclusivity of topics from journal articles.

No.	Topic Label	Semantic Coherence	Exclusivity
1	Logistics and International Growth	−160.05	11.83
2	Environmental Management and System	−170.77	11.54
3	Team, Role, and Manager	−155.52	11.81
4	Information Management and Organization	−85.94	11.13
5	Public Record Information	−134.78	11.08
6	Performance Quality Management	−99.56	11.72
7	Social Communication and Trust	−172.51	11.54
8	Construction Project Management	−122.69	11.56
9	Knowledge, Innovation, and Knowledge management	−123.60	11.57
10	Information System and Database	−131.45	11.15
11	Models and Systems for Decisions	−120.05	11.46
12	Information, Cost, Market	−156.93	11.63
13	Product and Enterprise Management	−125.70	11.62
14	User Adoption Models	−180.59	11.56
15	Data and Network Management	−167.50	11.20
16	Research, Management and Risks	−106.64	11.72

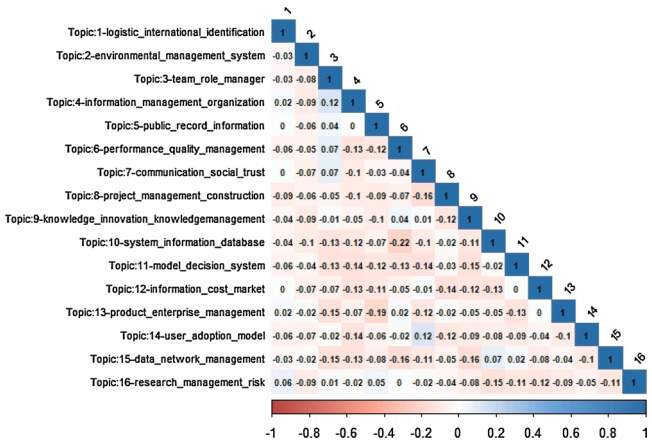


Fig. 19. Correlations among topics (only journal articles).

International Growth', 'Team, Role, and Manager', 'Information Management and Organization', 'Models and Systems for Decisions', and Topic 12-'Information, Cost, Market'.

4.3. Intersections between topics

The common themes and topics across all analyses are data management, knowledge management, supply chain management, information security management, environmental management, project management, service management, mobile and web management, healthcare management, and technology management. These themes and topics are well connected to IM domain because previous studies have consistently associated different aspects of IM with these themes like supply chain (Pereira, 2009), knowledge management (Costa, Soares, & De Sousa, 2016; Karim & Hussein, 2008; Larsen & Olaisen, 2013; Mao, Liu, Zhang, & Deng, 2016; Kruger & Johnson, 2010), information security management (Bang, Lee, Bae, & Ahn, 2012; Ding et al., 2014; Ogiela & Ogiela, 2018; Silva, De Gusmão, Poleto, Silva, & Costa, 2014; Soomro, Shah, & Ahmed, 2016), mobile, cloud and web management (Li, Qian, You, & Lu, 2014; Pinho, Franco, & Mendes, 2018; Saparova, Kibaru, & Bašić, 2013; Shiau, Dwivedi, & Lai, 2018; Yang, Wang, Gan, & Lin, 2013), project management (Garwood & Poole, 2018), operation management (Albergaria & Chiappetta Jabbour, 2019), healthcare (Foshay & Kuziemsky, 2014; Gao & Sunyaev, 2019; Hossain, Quaresma, & Rahman, 2019; Karaca, Moonis, Zhang, & Gezgez, 2019; Zheng et al., 2006), environmental management (Wang, Chen, & Benitez-Amado, 2015; Yang, Li, & Kang, 2020), and service management (Ali, Warren, & Mathiassen, 2017; Gibb, Buchanan, & Shah, 2006; Iden & Eikebrokk, 2013). Fig. 25 graphically represents the intersection between topics extracted from articles and conference papers.

It is evident from Fig. 25 that knowledge management, network management, project management, and environment management are common topics across the articles and conference papers. The findings of the study confirm that the IM domain is an evergreen field of research that attracts a growing number of scholars and researchers every year. The research landscape of the IM domain has evolved and witnessed a plethora of advancements during the selected time period of this study, specifically from 2000 to 2019. Thus, this domain has become more imperative and prevalent in recent years. This is also evident from the number of research documents published during the period 2000−2019. Perhaps the most notable change in the topic prevalences over the years from 1970 to 2019 can be witnessed during 2010, where several new themes emerged or became very popular suddenly. Interestingly, a few latent topics which are semantically similar to popular research themes in this era after 2010 like 'Social Media', 'Social Communication', 'Big Data', 'E-Commerce' and 'Knowledge Management' have emerged because researchers have relatively started to focus more on these themes.

4.4. A conceptual framework for information management

The intellectual and conceptual structure of the IM domain highlighting the knowledge, research, and practice is proposed by a four-level conceptual framework (Rowley, 1998). This framework tries to develop a complete understanding of IM at individual, system, context, and environment levels. The first two levels, information retrieval and information systems explain information processing and management by users and systems, which the author describes as *sub-discipline microinformatics*. The next two levels, information contexts and information environments, explain IM in a society, organization, and environment, which is described as *sub-discipline macroinformatics*. Optimal IM entails a careful focus at all levels, including individual, systems, contextual and societal levels. This study has adapted the same four levels framework (Rowley, 1998) to visually classify the extracted topics from conference papers and articles. The key themes emerging from the extant research are summarized in the proposed conceptual framework presented in Fig. 26. The conceptual framework puts forward the latent topics organized into individual, system, context, and environment levels. Consistent with Fig. 25, the topics in blue color are extracted from research articles, and topics in red are from conference papers.

Further, this study has also categorized the extracted topics as per the perspectives based framework proposed by Detlor (2010). The organizational, library, and personal perspectives of IM are presented in this seminal work. These perspectives use a process orientation to discuss the management of different processes that encompass the entire information lifecycle, including creation, acquisition, organization, storage, dissemination, and utilization of information. The latent topics discovered from IM documents are presented in Table 13, which classifies the topics as per organizational, library, and personal perspectives. It is evident that the organizational perspective predominantly covers a large number of topics related to IM, which is consistent with the findings by Detlor (2010). The personal perspective deals with the aspects related to personal IM, which has not gained more attention and importance in the domain. However, Detlor (2010) proposed that in the era of personal information processing devices like personal computers, the personal perspective may attract more researchers and scholars to study the personal factors associated with IM.

4.5. Implications for research

The traditional methods like bibliometrics, systematic literature

Table 11
Extracted topics and intuitive meanings-journal articles.

No.	Topic Label	Proportion	Top-10 Words	FREX	Lift
1	Logistics and International Growth	0.016	Logistic, International, Growth, Identification, Journal, Academic, RFID, Information Management, Growth, Technology	Logistic, International, Journal, RFID, Overload, Radio, Port, Identification, Reverse, European	Auto, Overload, Sub-disciplines, Tumor, Bibliometrics, ECIS, Edition, Port, Forwarder, Harbor
2	Environmental Management and System	0.034	Environmental, Management, System, Energy, Waste, Sustainability, ERP, Resource, Sustainable, Efficiency	Environmental, Waste, Water, Green, Carbon, Land, Emission, Ga, Oil, Recycle	Envelopment, Abiotic, Accounting, Acidification, Agronomic, Agronomy, Anaerobic, Aquatic, Archetypal, Beijing-Tianjin-Hebei
3	Team, Role, and Manager	0.036	Team, Role, Manager, Report, Employee, Work, Professional, Account, Development, Culture	Ethical, Executive, ISD, Employee, Team, Career, Chief, Programme, Job, Report	Career, Chief, E-marketplace, IPRS, Accountant, Appointment, CEO, CIO, CFO, Communal
4	Information Management and Organization	0.128	Information, Management, Organization, Information System, Business, Change, Technology, Plan, Information Management, System	Intelligence, Change, Organization, Crisis, Strategic planning, Science, Plan, Audit, Information technology, Reengineer	Await, Budgets spending, Catastrophic, Challenger, Chronology, Comprehensibility, E-strategic, Gartner, Genesis, HRD
5	Public Record Information	0.058	Public, Record, Information, Health, Healthcare, Management, Policy, Hospital, Government, Information management	Healthcare, Hospital, Medical, Patient, Emergency, Disaster, Clinical, Nurse, Record, Informatics	England, Silk, Adult, Agency's, AHA, Amendment, Antibiotic, Anti-infective, Antimicrobial, Arabia
6	Performance Quality Management	0.080	Performance, Quality, Management, Study, Factor, Firm, Model, Organizational, Relationship, Resource	Performance, Quality, Success, Alignment, Measurement, TQM, Capability, Measure, Questionnaire, Structural equation modeling	Baldrige, Cronbach's, CSFS, Depreciate, EFQM, Emirate, Fifty, Formality, Gulf, Insignificantly
7	Social Communication and Trust	0.045	Communication, Social, Trust, Online, Study, Theory, Community, Governance, Medium, ICT	SMEs, Social-media, Cultural, Trust, Social-network, Cross-cultural, Hotel, ICT, Email, Social	Abductive, Aggressiveness, Benevolence, Biographical, CBT, Collectivism, Collectivisms, Colocation, Confucian, Conscientiousness
8	Construction Project Management	0.075	Project, Management, Construction, Design, Engineer, Information, Process, System, Model, Software	Project, Construction, BIM, Contractor, Maintenance, Concurrent, ASCE, Engineer, Schedule, Workflow	Adapter, Adaption, plant manufacturing, Bug, Build time, CAE, Cavern, CICS, CII, Company-specific
9	Knowledge, Innovation, and Knowledge Management	0.068	Knowledge, Innovation, Knowledge management, Learn, Transfer, Knowledge sharing, Firm, Capital, Technology, Organizational	Knowledge management, Intellectual, Patent, Tacit, Innovation, Knowledge sharing, Knowledge, Absorptive, Transfer, Learn	Acquirers, appropriability, Catalyze, Coevolution, Commercialization, Commercialization, Conduit, Coopetition, Cross-fertilization, Depiction
10	Information System and Database	0.070	System, Information, Database, Computer, Data, Document, Web, Information management, Digital, User	Semantic, Ontology, Xml, Multimedia, Schema, Language, Database, Document, Query, Retrieval	Calendar, Compression, Corpus, Federate, Multilingual, Ontology, Www, Acid, Alias, Ancient
11	Models and Systems for Decisions	0.079	Model, Decision, System, Analysis, Method, Decision-making, Support, Process, Problem, Evaluation	Fuzzy, DSS, Decision, Classification, Datamining, Algorithm, Heuristic, Decision-making, Estimation, Neural	ABC, Apriori, Bowman, BOW, Chance constrained, DSS, Comprehensible, Discrete event, DST, DTA
12	Information, Cost, Market	0.060	Information, Cost, Market, Share, Supply chain, Demand, Inventory, Outsource, Control, Price	Price, Tourism, Auction, Retailer, Buyer, Inventory, Outsource, Game theory, Retail, Principal agent	Encroachment, Reorder, Abate, Advertiser, Alpine, Alternately, Arbitration, Arm's length, Auction, Backorder
13	Product and Enterprise Management	0.069	Product, Enterprise, Management, Service, Manufacture, Industry, Supply chain, Integration, Customer, Process	CRM, Manufacture, Product, Enterprise, NPD, Textile, Integration, Customer, Lean, Flexibility	SCM, Auto, Blueprinting, Compustat, CRM, Cross company, Cross sell, Customer centric, Customer facing, Customer oriented
14	User Adoption Models	0.046	User, Adoption, Model, Perceive, Service, Study, Consumer, Mobile, Satisfaction, Behavior	Intention, Service quality, Technology acceptance, Bank, Adoption, Self-efficacy, Satisfaction, Attitude, Continuance, Perceive	M-commerce, Affinity, AINSI, Altruism, Amos, Antivirus, Approaches, Aux, Avec, Banner
15	Data and Network Management	0.056	Data, Network, Management, System, Big data, Monitor, Control, Information management, Operation, Smart	Traffic, Sensor, Data warehouse, IBM, Pavement, Smart, Realtime, Analytics, Data warehouses, Big data	Airborne, Alberta, Analog, Anthropometry, Apache, ATC, ATM, Attacker, Boot, Bottle
16	Research, Management and Risks	0.082	Research, Management, Risk, Study, Framework, Finding, Limit, Security, Approach, Analysis	Information security, Originality value, Design methodology approach, Risk, Limitation, implications, Review, Practical, Limit, Security, Qualitative	CIA, Constructionist, E-science, NIST, Triangulate, Information security, Isms, Configurability, Cyberspace, Discontinuation

review, meta-analysis, and research thematic analysis are established methods from the last several decades, and many researchers have used these methods for integrating the vast and fragmented research into a comprehensive, systematic, objective, and transparent representation. This study proposes a new methodology to assimilate the fragmented knowledge in order to discover interesting insights in an unsupervised way, which can be very useful for the current and potential researchers of the IM domain. Understanding and reporting the evolution of the overall intellectual structure of the IM research for the last fifty years may add a significant value to the existing body of knowledge, and such study would also be potentially interesting for academicians and practitioners in this field.

This study may serve as a gateway to IM for readers and aspiring researchers in this field. The outcomes of this study offer a comprehensive snapshot of IM knowledge structure to its potential contributors. The study offers several contributions to the IM domain by summarizing the important publications and the most influential authors as well as by visualizing the evolution of latent themes and topics over fifty years. Further, the insights and conclusions presented in the study may be significant for editorial boards of various journals for formulating future policies related to the inclusion of new and emerging themes in special issues as well as framing strategies for further development of IM knowledge. Another contribution of this study is the novel hybrid methodology, which offers an alternative to traditional bibliometric analyses based methods. This study proposes sophisticated text analytics based approach utilizing a probabilistic generative model based on

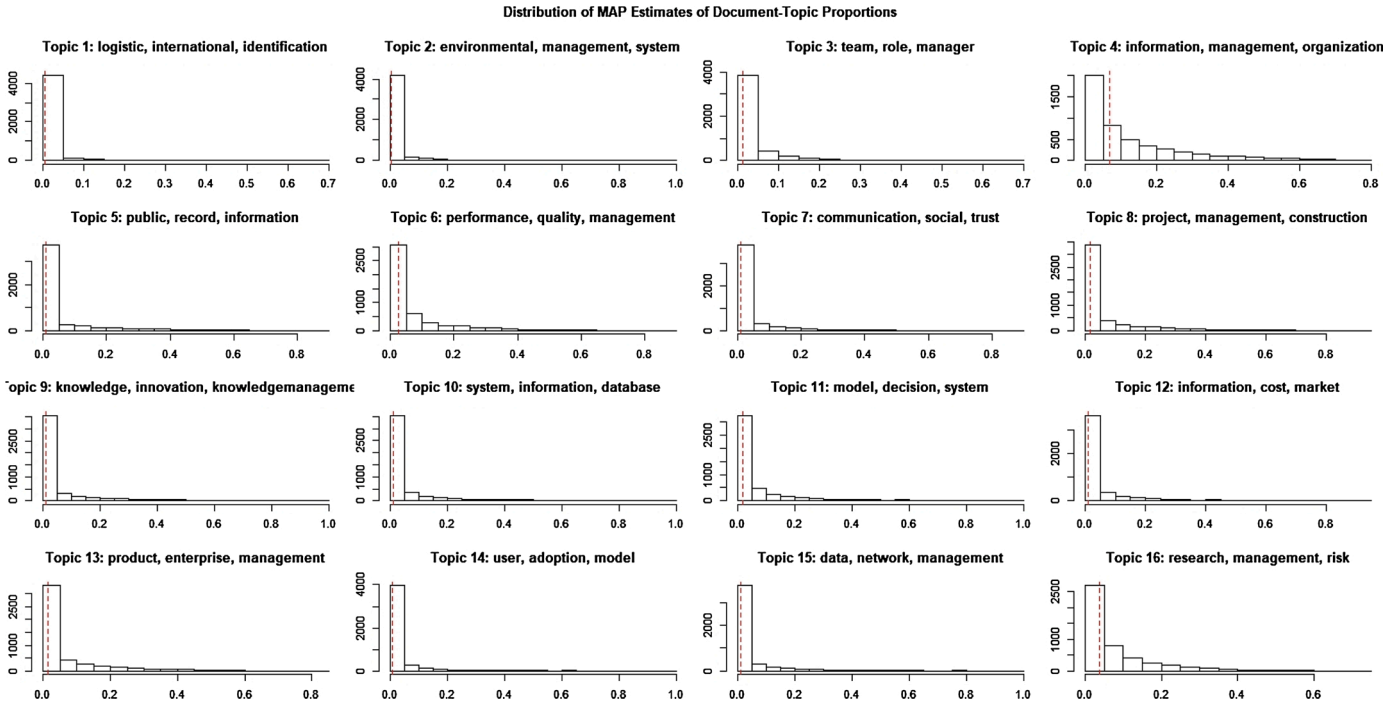


Fig. 20. Topic proportions within documents-journal articles.

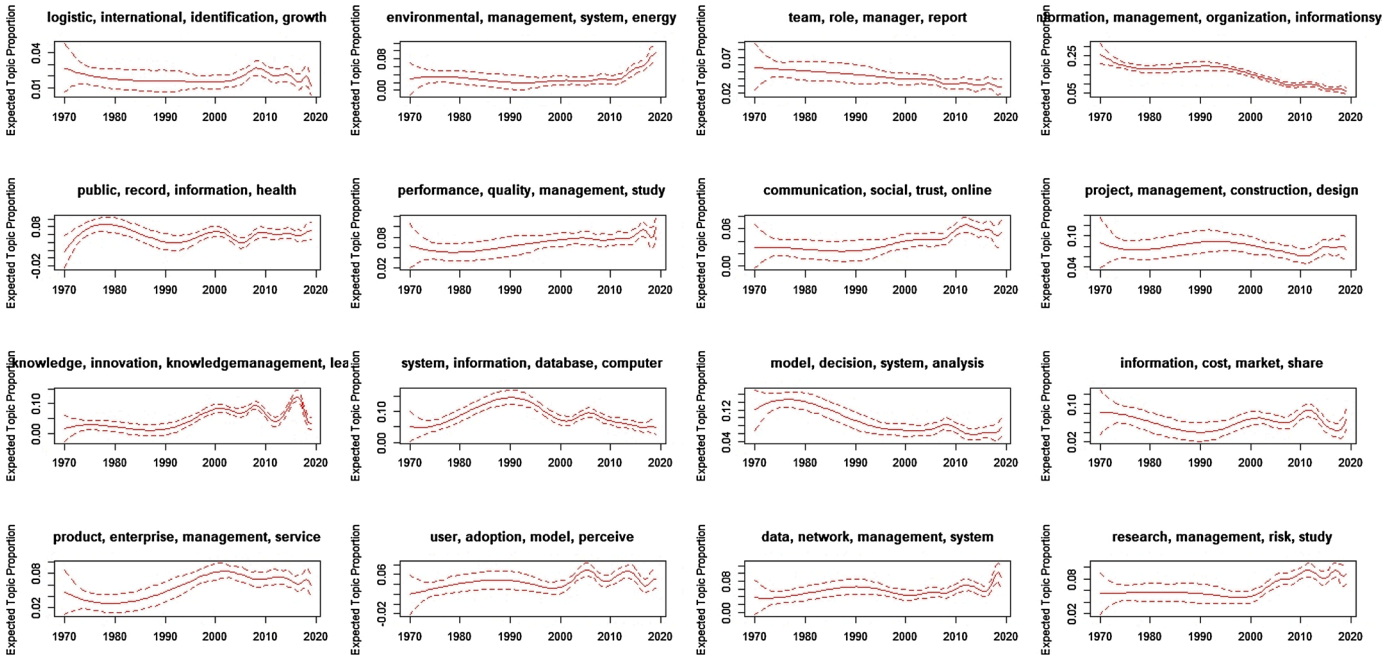


Fig. 21. Covariate effect estimation of the topic prevalences over years (journal articles).

structural topic modeling. The STM based topic modeling emphasizes on identifying, analyzing, interpreting, summarizing, and visualizing the key themes which are latent in the plethora of text documents. Furthermore, this study also identifies emerging themes and topics. Hence, this study goes beyond previous literature by revealing the latent themes where future work will be interesting to undertake by potential researchers and practitioners of this domain. The latent themes that indicate a rising trend over the years can be further explored by the potential researchers in this domain.

4.6. *Limitations and directions for future research*

This study has utilized the peer-reviewed research data collected from Scopus, which is a major limitation of this work. Although, many previous studies have used Scopus, assuming it the largest source of research abstract and citation database (Donthu et al., 2020; Singh et al., 2020). However, this study has not compared Scopus with other sources of equivalent coverage and quality like Web of Science and Google Scholar. This study supports the view that such a comparison could be out of scope and context as per the objectives of the current study. The citation details for research articles, authors, journals, institutions, and

Table 12
Thematic evolution of information management domain from 1970 to 2019 (top-10 keywords and frequencies).

Document Type	1970–1979	1980–1989	1990–1999	2000–2009	2010–2019
All Documents	information management (4), data handling (3), administrative data processing (1), aptitude tests (1), aura of credibility (1), corporate management (1), career paths (1), computer programmers (1), systems programming (1), computers (1)	information management (37), management (12), information systems (11), decision support systems (7), management information systems (7), office automation (7), planning (6), information resource management (5), information science (5), computer applications (4)	information management (483), information technology (99), management information systems (81), strategic planning (75), database systems (72), decision support systems (57), information systems (50), information retrieval systems (47), mathematical models (46), computer software (43)	information management (3696), innovation (656), information technology (598), industrial engineering (595), knowledge management (415), societies and institutions (382), industrial management (353), information systems (320), decision making (304), electronic commerce (282)	information management (13029), innovation (3448), sustainable development (3207), economics (2862), regional planning (2032), industrial engineering (1454), competition (1425), economic and social effects (1213), information systems (1209), ecommerce (1110)
Only Conference Papers	information management (4), data handling (3), administrative data processing (1), information services (1), career paths (1), computer programmers (1), computer systems programming (1), corporate management (1), data center operations (1), data processing departments (1)	information management (3), management information systems (3), information systems (2), information use (2), office automation (2), access control (1), production control (1), computer software - selection (1), conceptual frameworks (1), document-processing (1)	information management (134), database systems (31), strategic planning (28), manufacture (22), information technology (18), inventory control (18), marketing (18), information retrieval systems (17), quality assurance (16), management information systems (13)	information management (2094), innovation (604), industrial engineering (585), information technology (243), knowledge management (172), electronic commerce (167), competition (157), information systems (147), industrial management (134), societies and institutions (133)	information management (9444), sustainable development (3105), economics (2716), innovation (2677), regional planning (2011), competition (1218), economic and social effects (1156), ecommerce (961), information systems (819), industrial engineering (816)
Only Articles	health care - management (1), computers (1), data base systems (1), data base systems (1), data processing - hospital applications (1), data processing, business - legislation (1), data storage devices (1), decision points (1), detrimental reliance (1), erroneous computer-output (1)	information management (34), management (12), information systems (9), decision support systems (7), planning (6), information resource management (5), information science (5), office automation (5), computer applications (4), data processing (4)	information management (337), information technology (80), management information systems (67), decision support systems (51), strategic planning (46), mathematical models (42), database systems (41), information systems (39), decision making (34), societies and institutions (34)	information management (1349), information technology (299), knowledge management (225), societies and institutions (216), industrial management (188), decision making (156), strategic planning (156), information systems (146), mathematical models (142), project management (136)	information management (1734), knowledge management (321), information systems (221), decision making (186), project management (150), design/methodology/ approach (144), supply chain management (116), human resource management (105), surveys (95), innovation (89)

countries used in this study are based on Scopus data at a point of time only, which may evolve over time and differ if any other data source is considered. Further, the search protocol used in this study has restricted the bibliometric data to the Business, Management, and Accounting subject area only. For future work, we suggest that researchers might explore the current methodology on other databases with other subject areas (e.g. social sciences, decision sciences, computer science etc.). The results of any bibliometrics-based study do change over time, which makes it difficult to reproduce the same results in the future. Moreover, the domains like information management, information systems, and information technology are dynamic, and they witness the emergence of new mainstream themes every year, which merits additional studies performing static prospective and retrospective analysis at fixed intervals.

The mainstream and emerging themes visualized in this study can be explored further by future studies in the IM domain. This study has visually presented that themes and topics like knowledge management, supply chain management, information security management, environmental management, project management, service management, mobile and web management, healthcare applications management, and technology management are emerging from the IM domain, reconfirming the truly multidisciplinary nature of this domain. Interestingly, the growth of multidisciplinary research related to IM seems to be quite focused on these topics and themes in nature. This conclusion is based on the analysis of the topic prevalence from 1970 to 2019, which revealed that there are cumulative trends in these topics. This offers a guide and roadmap for future research endeavors, which can be potentially useful for readers and aspiring researchers in the domain of information management.

5. Conclusion

IM has become a key organizational activity that manages the people, processes, and technologies related to the generation, acquisition, storage, processing, dissemination, and disposition of information. Information resource management started in the 1970s embracing all the generic activities related to data management with the proliferation of information technology, and soon it became a significant factor behind organizational survival and success. However, over the years, the IM domain has witnessed exponential growth with a convergence of various themes, domains, and disciplines. Extraction of these themes and understanding the chronological evolution of these research themes are the main objectives of this study. This study has summarized and visualized the evolution of the IM domain by identifying key publications, authors, and keywords and exploring how specific themes have evolved over time. This study has conducted structural topic modeling to extract the latent topics and the hidden thematic domain knowledge structure from the vast literature related to the IM domain and visualized the covariate effect estimation of the topic prevalences from all the research documents published over the years starting from 1970 to 2019. The STM

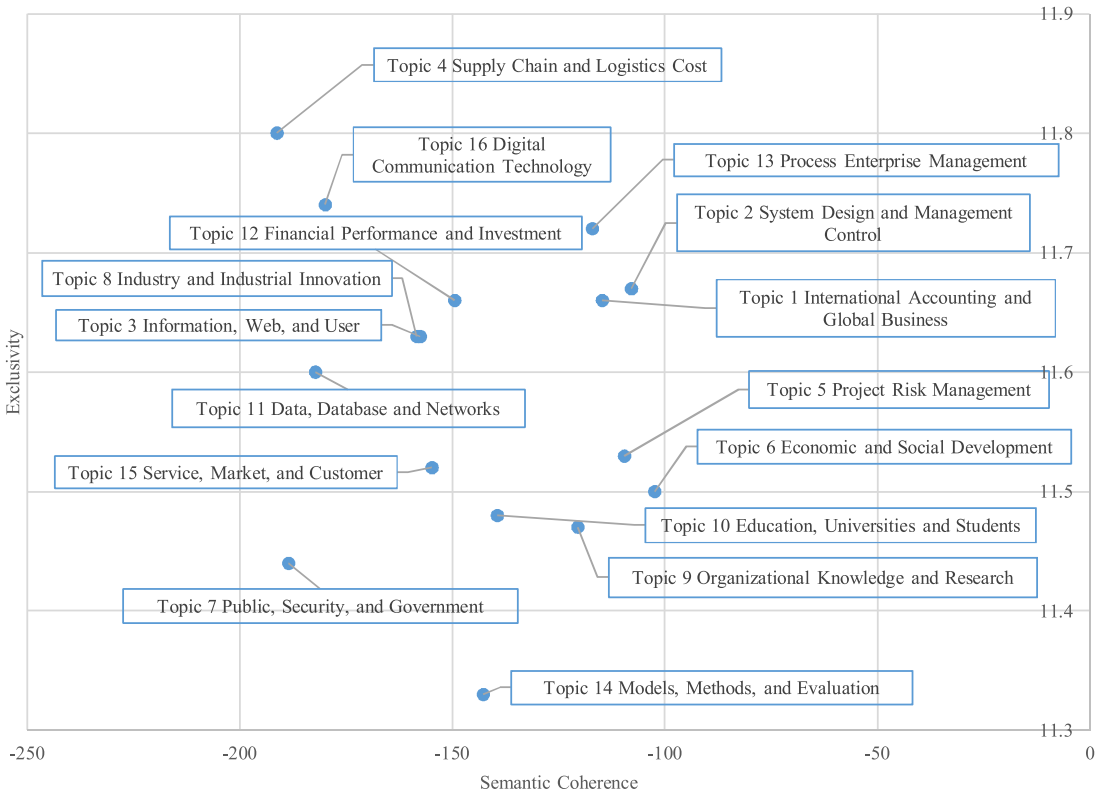


Fig. 22. Average semantic coherence and exclusivity to show topic quality (all papers).

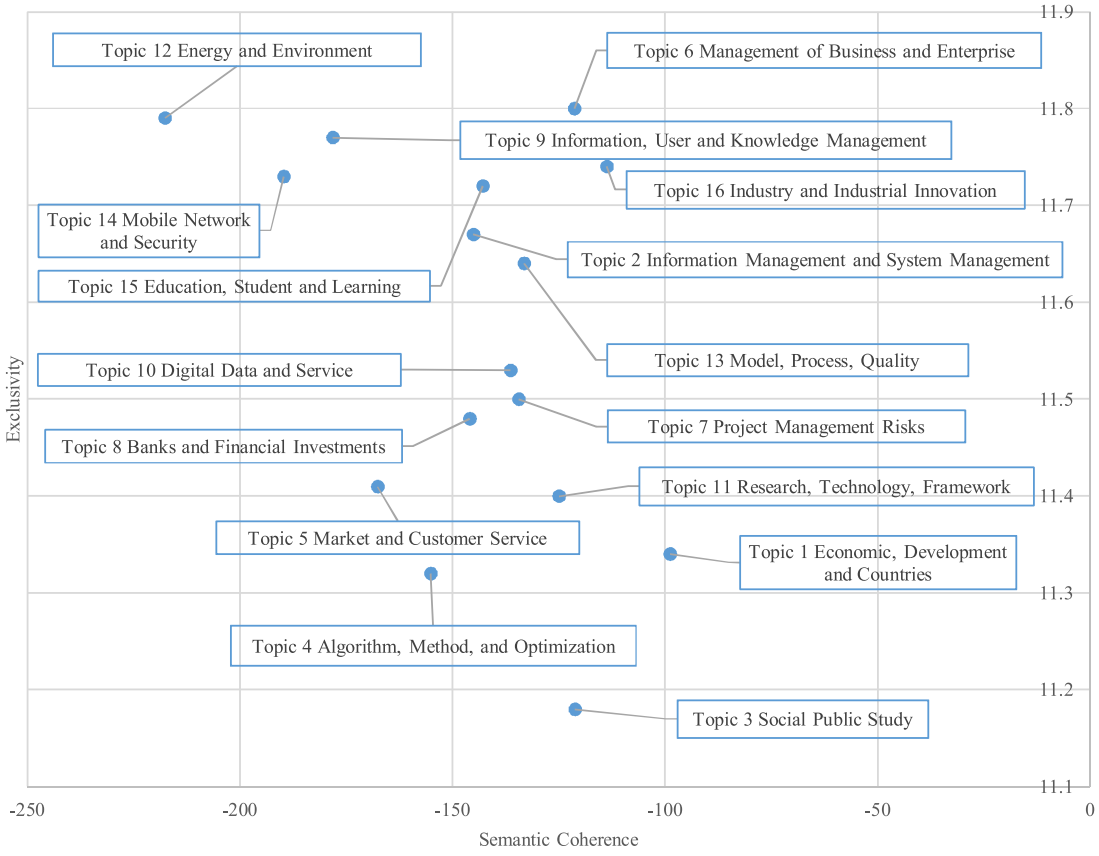


Fig. 23. Semantic coherence and exclusivity to show topic quality (conference papers).

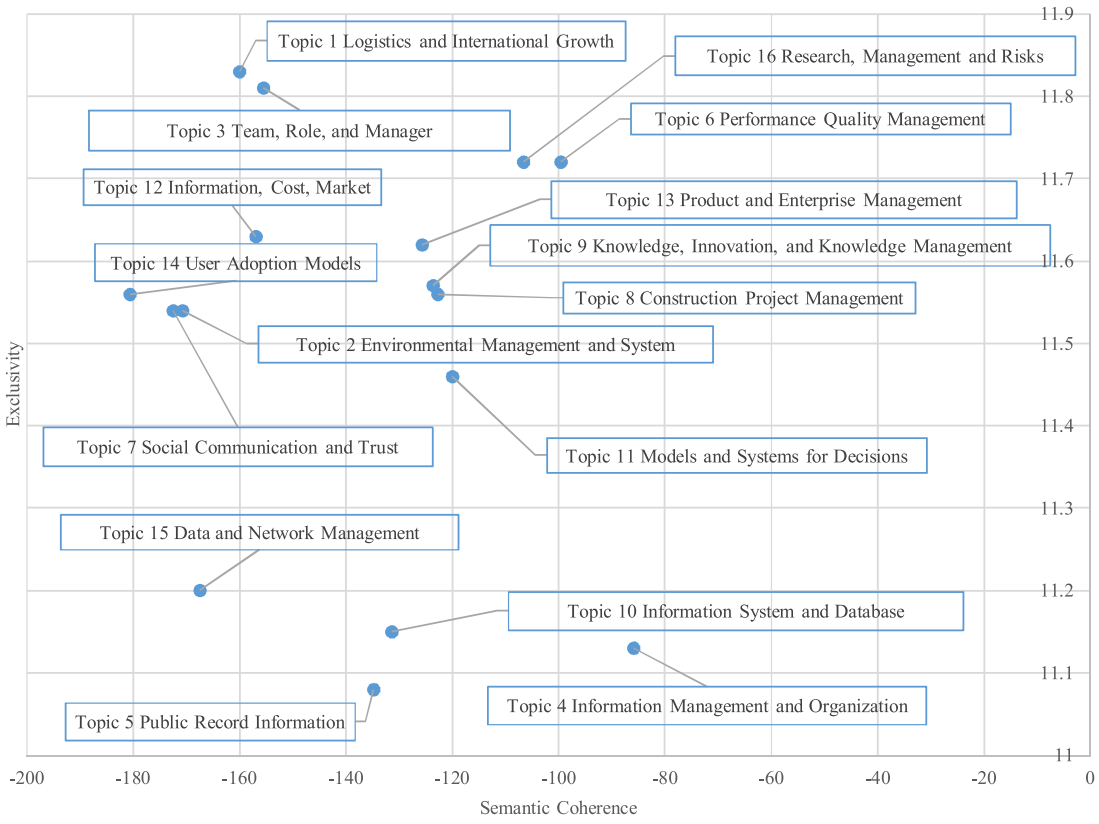


Fig. 24. Average semantic coherence and exclusivity to show topic quality (articles).

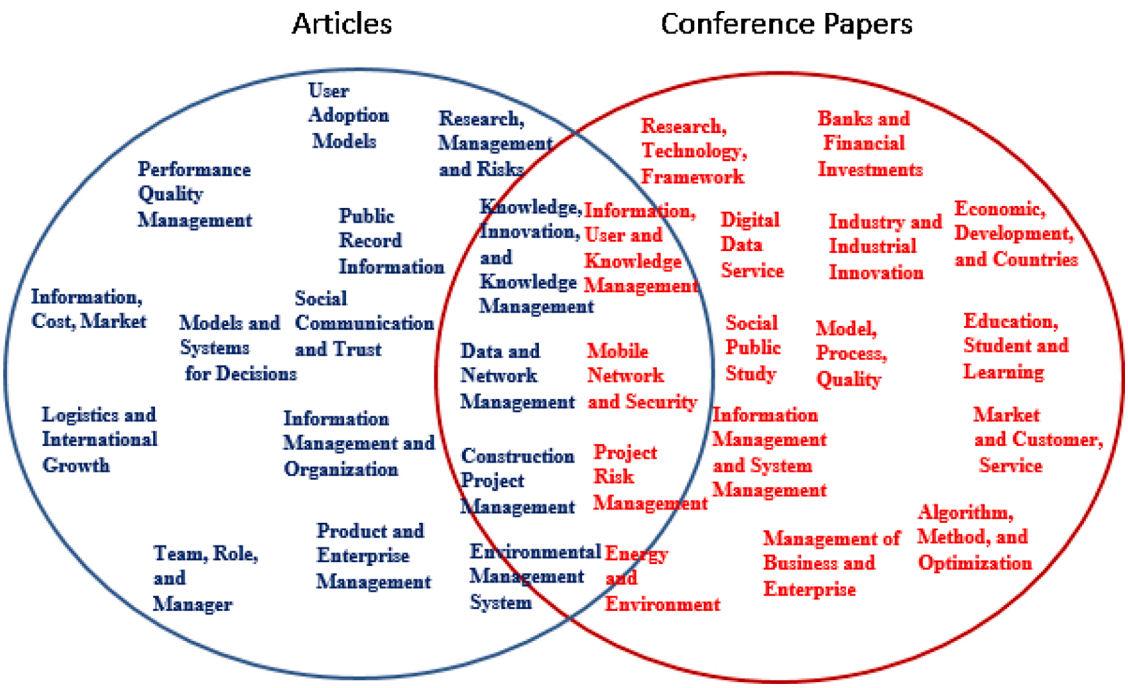


Fig. 25. Intersections between topics extracted from articles and conference papers.

bases topic modeling is performed for all the documents first, and then separate analyses are conducted on conference papers and journals' articles. The results have shown that the most common themes across all the analyses are related to data management, knowledge management, supply chain management, information security management, environmental management, project management, service management, mobile and web management, healthcare application management, and technology management which reconfirms that IM is a multi-disciplinary domain in real sense. This study has highlighted that STM is an advanced text analytics based method which enables to extract more interesting insights as latent topics as well as model the hidden thematic structure of the domain knowledge.

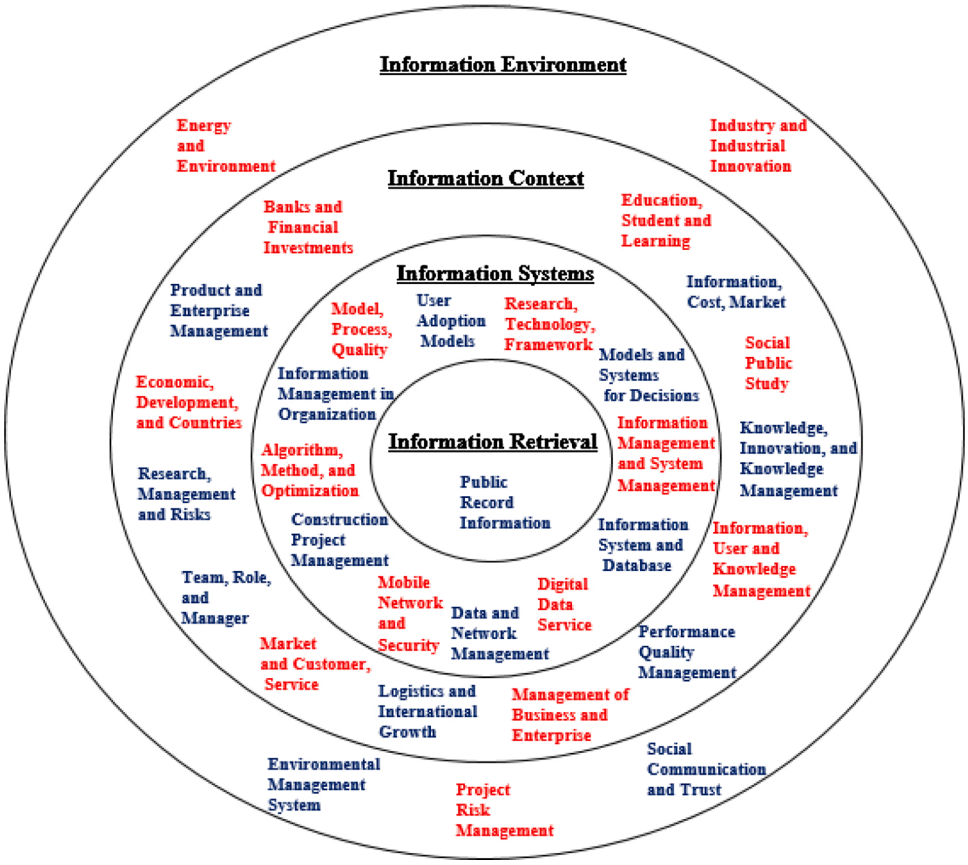


Fig. 26. A proposed framework for information management (Adapted from Rowley, 1998).

Table 13
Classifying topics as per organizational, library, and personal perspectives (Adapted from Detlor, 2010).

	Organizational Perspective	Library Perspective	Personal Perspective
All Documents	International Accounting and Global Business Supply Chain and Logistics Cost Project Risk Management Economic and Social Development Industry and Industrial Innovation Organizational Knowledge and Research Financial Performance and Investment Process Enterprise Management	System Design and Management Control Data, Database and Networks Models, Methods, and Evaluation Digital Communication Technology	Public Security and Government Education, Universities and Students Service, Market, and Customer Information, Web, and User
Conference Papers	Economic, Development, and Countries Social Public Study Market and Customer, Service Management of Business and Enterprise Project Risk Management Banks and Financial Investments Energy and Environment Model, Process, Quality Mobile Network and Security Industry and Industrial Innovation	Information Management and System Management Algorithm, Method, and Optimization Digital Data and Service Research, Technology, Framework	Education, Student and Learning Information, User and Knowledge Management
Articles	Logistics and International Growth Environmental Management System Information Management and Organization Performance Quality Management Social Communication and Trust Construction Project Management Models and Systems for Decisions Information, Cost, Market Product and Enterprise Management Research, Management and Risks	Public Record Information Knowledge, Innovation, and Knowledge management Information System and Database Data and Network Management	Team, Role, and Manager User Adoption Models

CRediT authorship contribution statement

Anuj Sharma: Conceptualization, Methodology, Data curation, Writing - original draft. **Nripendra P. Rana:** Conceptualization, Methodology, Writing - review & editing, Supervision. **Robin Nunkoo:** Visualization, Validation, Writing- review & editing, Supervision.

References

Abedin, B., Jafarzadeh, H., & Olszak, C. M. (2020). Thirty six years of information systems management: A bibliometric and thematic analysis. *Information Systems Management*, 1–14. https://doi.org/10.1080/10580530.2020.1781987.

Acedo, F. J., & Casillas, J. C. (2005). Current paradigms in the international management field: An author co-citation analysis. *International Business Review, 14*(5), 619–639. https://doi.org/10.1016/j.ibusrev.2005.05.003.

Adelman, A., & Kemp, A. J. (1970). Space station information management. *IEEE Transactions on Aerospace and Electronic Systems, AES-6*(6), 746–763. https://doi.org/10.1109/TAES.1970.310156.

Ahmad, I., Ahmed, G., Shah, S. A. A., & Ahmed, E. (2020). A decade of big data literature: Analysis of trends in light of bibliometrics. *The Journal of Supercomputing, 76*(5), 3555–3571. https://doi.org/10.1007/s11227-018-2714-x.

Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly, 25*(1), 107–136. https://doi.org/10.2307/3250961.

Albergaria, M., & Chiappetta Jabbour, C. J. (2019). The role of big data analytics capabilities (BDAC) in understanding the challenges of service information and operations management in the sharing economy: Evidence of peer effects in libraries. *International Journal of Information Management*. https://doi.org/10.1016/j.ijinfomgt.2019.10.008.

Ali, A., Warren, D., & Mathiassen, L. (2017). Cloud-based business services innovation: A risk management model. *International Journal of Information Management, 37*(6), 639–649. https://doi.org/10.1016/j.ijinfomgt.2017.05.008.

Alsudairi, M., & Dwivedi, Y. K. (2010). A multi-disciplinary profile of IS/IT outsourcing research. *Journal of Enterprise Information Management, 23*(2), 215–258. https://doi.org/10.1108/17410391021019787.

Anderton, R. H. (1986). Post-graduate education for information management in the UK: New developments at Lancaster. *International Journal of Information Management, 6*(4), 247–249. https://doi.org/10.1016/0268-4012(86)90025-3.

Ardito, L., Scuotto, V., Del Giudice, M., & Petruzzelli, A. M. (2019). A bibliometric analysis of research on Big Data analytics for business and management. *Management Decision, 57*(8), 1993–2009. https://doi.org/10.1108/MD-07-2018-0754.

Avison, D. E., Dwivedi, Y. K., Fitzgerald, G., & Powell, P. (2008). The beginnings of a new era: Time to reflect on 17 years of the ISJ. *Information Systems Journal, 18*(1), 5–21. https://doi.org/10.1111/j.1365-2575.2007.00282.x.

Bang, Y., Lee, D.-J., Bae, Y.-S., & Ahn, J.-H. (2012). Improving information security management: An analysis of ID-password usage and a new login vulnerability measure. *International Journal of Information Management, 32*(5), 409–418. https://doi.org/10.1016/j.ijinfomgt.2012.01.001.

Bate, V. D., & Gedam, J. (2019). SRELS journal of information management: A bibliometric study. *Library Philosophy and Practice (e-journal)*, 3035.

Beimesch, B. K. B. (1982). Organizational structure and its impact upon information management. *Information Services & Use, 2*(1), 35–40. https://doi.org/10.3233/ISU-1982-2104.

Best, D. P. (1988). The future of information management. *International Journal of Information Management, 8*(1), 13–24. https://doi.org/10.1016/0268-4012(88)90004-7.

Blazewicz, J. F. (1983). Information management tool for system designers. *Electronic Systems Technology and Design/Computer Design's, 22*(4). Retrieved from Scopus.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77–84. https://doi.org/10.1145/2133806.2133826.

Blei, D. M., & Lafferty, J. D. (2005). *Correlated topic models* (pp. 147–154). Retrieved from Scopus.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*(4–5), 993–1022. Retrieved from Scopus.

Broadus, R. N. (1987). Toward a definition of "bibliometrics". *Scientometrics, 12*(5), 373–379. https://doi.org/10.1007/BF02016680.

Brussaard, B. K. (1983). Electronic information management for administrative government. *Interdisciplinary Science Reviews, 8*(1), 56–64. https://doi.org/10.1179/isr.1983.8.1.56.

Brussaard, B. K. (1988). Information resource management in the public sector. *Information & Management, 15*(2), 85–92. https://doi.org/10.1016/0378-7206(88)90039-0.

Callon, M., Courtial, J.-P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Information (International Social Science Council), 22*(2), 191–235. https://doi.org/10.1177/053901883022002003.

Carter, L. F. (1983). Electronic information management in the social sciences. *Interdisciplinary Science Reviews, 8*(1), 76–84. https://doi.org/10.1179/isr.1983.8.1.76.

Cheng, T. C. E. (1987). Toward a policy framework for business information resources management. *Industrial Management & Data Systems, 87*(1–2), 5–8. https://doi.org/10.1108/eb057463.

Cobo, M. J., Martínez, M. A., Gutiérrez-Salcedo, M., Fujita, H., & Herrera-Viedma, E. (2015). 25years at knowledge-based systems: A bibliometric analysis. *Knowledge-based Systems, 80*, 3–13. https://doi.org/10.1016/j.knosys.2014.12.035.

Costa, E., Soares, A. L., & De Sousa, J. P. (2016). Information, knowledge and collaboration management in the internationalisation of SMEs: A systematic literature review. *International Journal of Information Management, 36*(4), 557–569. https://doi.org/10.1016/j.ijinfomgt.2016.03.007.

Davenport, T. H. (1998). Putting the enterprise into the enterprise system. *Harvard Business Review, 76*(4), 121–131.

Davenport, L., & Cronin, B. (1988). Strategic information management: Forging the value chain. *International Journal of Information Management, 8*(1), 25–34. https://doi.org/10.1016/0268-4012(88)90005-9.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science, 41*(6), 391–407. https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASI1>3.0.CO;2-9.

Detlor, B. (2010). Information management. *International Journal of Information Management, 30*(2), 103–108. https://doi.org/10.1016/j.ijinfomgt.2009.12.001.

Ding, J.-H., Chien, R., Hung, S.-H., Lin, Y.-L., Kuo, C.-Y., Hsu, C.-H., … Chung, Y.-C. (2014). A framework of cloud-based virtual phones for secure intelligent information management. *International Journal of Information Management, 34*(3), 329–335. https://doi.org/10.1016/j.ijinfomgt.2013.11.006.

Donthu, N., Kumar, S., Pandey, N., & Gupta, P. (2021). Forty years of the International Journal of Information Management: A bibliometric analysis. *International Journal of Information Management, 57*, Article 102307. https://doi.org/10.1016/j.ijinfomgt.2020.102307.

Donthu, N., Kumar, S., & Pattnaik, D. (2020). Forty-five years of Journal of Business Research: A bibliometric analysis. *Journal of Business Research, 109*, 1–14. https://doi.org/10.1016/j.jbusres.2019.10.039.

Dwivedi, Y. K., & Kuljis, J. (2008). Profile of IS research published in the European journal of information systems. *European Journal of Information Systems, 17*(6), 678–693. https://doi.org/10.1057/ejis.2008.57.

Dwivedi, Y. K., & Mustafee, N. (2010). Profiling research published in the journal of enterprise information management. *Journal of Enterprise Information Management, 23*(1), 8–26.

Dwivedi, Y. K., Lal, B., Mustafee, N., & Williams, M. D. (2008). Profiling a decade of information systems frontiers' research. *Information Systems Frontiers, 11*(1), 87. https://doi.org/10.1007/s10796-008-9147-7.

Ellis, D. (1986). Bibliographic essay. Information management and information work. *International Journal of Information Management, 6*(2), 115–116. https://doi.org/10.1016/0268-4012(86)90016-2.

Entsua-Mensah, C. (1996). Towards effective information management: A view from Ghana. *International Journal of Information Management, 16*(2), 149–156. https://doi.org/10.1016/0268-4012(95)00075-5.

Espejo, R. (1983). Strategies for information management. *Cybernetics and Systems, 14*(2–4), 315–341. https://doi.org/10.1080/01969728308927730.

Fetscherin, M., Voss, H., & Gugler, P. (2010). 30 Years of foreign direct investment to China: An interdisciplinary literature review. *International Business Review, 19*(3), 235–246. https://doi.org/10.1016/j.ibusrev.2009.12.002.

Foshay, N., & Kuziemsky, C. (2014). Towards an implementation framework for business intelligence in healthcare. *International Journal of Information Management, 34*(1), 20–27. https://doi.org/10.1016/j.ijinfomgt.2013.09.003.

Fuller-Love, N., & Cooper, J. (1996). Competition or co-operation? Strategic information management in the National Health Service: A case study of the Ceredigion NHS Trust. *International Journal of Information Management, 16*(3), 219–232. https://doi.org/10.1016/0268-4012(96)00006-0.

Gao, F., & Sunyaev, A. (2019). Context matters: A review of the determinant factors in the decision to adopt cloud computing in healthcare. *International Journal of Information Management, 48*, 120–138. https://doi.org/10.1016/j.ijinfomgt.2019.02.002.

Garwood, D. A., & Poole, A. H. (2018). Project management as information management in interdisciplinary research: "Lots of different pieces working together". *International Journal of Information Management, 41*, 14–22. https://doi.org/10.1016/j.ijinfomgt.2018.03.002.

Gibb, F., Buchanan, S., & Shah, S. (2006). An integrated approach to process and service management. *International Journal of Information Management, 26*(1), 44–58. https://doi.org/10.1016/j.ijinfomgt.2005.10.007.

Glänzel, W. (1996). A bibliometric approach to social sciences. National research performances in 6 selected social science areas, 1990–1992. *Scientometrics, 35*(3), 291–307. https://doi.org/10.1007/BF02016902.

González-Valiente, C. L., León Santos, M., Arencibia-Jorge, R., Noyons, E., & Costas, R. (2019). Mapping the evolution of intellectual structure in information management using author co-citation analysis. *Mobile Networks and Applications*. https://doi.org/10.1007/s11036-019-01231-9.

Gregor, S. (2006). The nature of theory in Information Systems. *MIS Quarterly, 30*(3), 611–642.

Hansen, M. T., Nohria, N., & Tierney, T. (1999). *What's your strategy for managing knowledge?*. Retrieved from https://www.hbs.edu/faculty/Pages/item.aspx?num=7313.

Herring, J. E. (1991). Information management—The convergence of professions. *International Journal of Information Management, 11*(2), 144–155. https://doi.org/10.1016/0268-4012(91)90005-W.

Hoard, J. E. (1989). Natural language processing and advanced information management. *Telematics and Informatics, 6*(3), 313–329. https://doi.org/10.1016/S0736-5853(89)80024-3.

Hofmann, T. (1999). *Probabilistic latent semantic indexing* (pp. 50–57). https://doi.org/10.1145/312624.312649.

Hossain, A., Quaresma, R., & Rahman, H. (2019). Investigating factors influencing the physicians' adoption of electronic health record (EHR) in healthcare system of Bangladesh: An empirical study. *International Journal of Information Management, 44*, 76–87. https://doi.org/10.1016/j.ijinfomgt.2018.09.016.

Hu, N., Zhang, T., Gao, B., & Bose, I. (2019). What do hotel customers complain about? Text analysis using structural topic model. *Tourism Management, 72*, 417–426. https://doi.org/10.1016/j.tourman.2019.01.002.

Huotari, M.-L. (1995). Strategic information management: A pilot study in a Finnish pharmaceutical company. *International Journal of Information Management, 15*(4), 295–302. https://doi.org/10.1016/0268-4012(95)00025-3.

Iden, J., & Eikebrokk, T. R. (2013). Implementing IT service management: A systematic literature review. *International Journal of Information Management, 33*(3), 512–523. https://doi.org/10.1016/j.ijinfomgt.2013.01.004.

Jackson, M. C. (1987). Systems strategies for information management in organizations which are not machines. *International Journal of Information Management, 7*(4), 187–195. https://doi.org/10.1016/0268-4012(87)90012-0.

Jellis, J. (1988). Information management: Meaning and learning. *International Journal of Information Management, 8*(1), 35–42. https://doi.org/10.1016/0268-4012(88)90006-0.

Karaca, Y., Moonis, M., Zhang, Y.-D., & Gezgez, C. (2019). Mobile cloud computing based stroke healthcare system. *International Journal of Information Management, 45*, 250–261. https://doi.org/10.1016/j.ijinfomgt.2018.09.012.

Karim, N. S. A., & Hussein, R. (2008). Managers' perception of information management and the role of information and knowledge managers: The Malaysian perspectives. *International Journal of Information Management, 28*(2), 114–127. https://doi.org/10.1016/j.ijinfomgt.2007.08.003.

Khanra, S., Dhir, A., & Mäntymäki, M. (2020). Big data analytics and enterprises: A bibliometric synthesis of the literature. *Enterprise Information Systems, 14*(6), 737–768. https://doi.org/10.1080/17517575.2020.1734241.

Kim, J., & McMillan, S. J. (2008). Evaluation of internet advertising research: A bibliometric analysis of citations from key sources. *Journal of Advertising, 37*(1), 99–112. https://doi.org/10.2753/JOA0091-3367370108.

Kim, J., Kang, S., & Lee, K. H. (2019). Evolution of digital marketing communication: Bibliometric analysis and network visualization from key articles. *Journal of Business Research*. https://doi.org/10.1016/j.jbusres.2019.09.043.

King, K. J., & Maryanski, F. J. (1983). Information management trends in office automation. *Proceedings of the IEEE, 71*(4), 519–528. https://doi.org/10.1109/PROC.1983.12623.

Kruger, C. J., & Johnson, R. D. (2010). Information management as an enabler of knowledge management maturity: A South African perspective. *International Journal of Information Management, 30*(1), 57–67. https://doi.org/10.1016/j.ijinfomgt.2009.06.007.

Kuhn, K. D. (2018). Using structural topic modeling to identify latent topics and trends in aviation incident reports. *Transportation Research Part C, Emerging Technologies, 87*, 105–122. https://doi.org/10.1016/j.trc.2017.12.018.

Laengle, S., Merigó, J. M., Miranda, J., Słowiński, R., Bomze, I., Borgonovo, E., … Teunter, R. (2017). Forty years of the European Journal of Operational Research: A bibliometric overview. *European Journal of Operational Research, 262*(3), 803–816. https://doi.org/10.1016/j.ejor.2017.04.027.

Laengle, S., Modak, N. M., Merigó, J. M., & Sotta, C. D. L. (2018). Thirty years of the international journal of computer integrated manufacturing: A bibliometric analysis. *International Journal of Computer Integrated Manufacturing, 31*(12), 1247–1268. https://doi.org/10.1080/0951192X.2018.1529434.

Larsen, T. J., & Olaisen, J. (2013). Innovating strategically in information and knowledge management: Applications of organizational behavior theory. *International Journal of Information Management, 33*(5), 764–774. https://doi.org/10.1016/j.ijinfomgt.2013.05.003.

Lavallee, W. L. (1983). The role of information management in nuclear plant operating experience review. *Progress in Nuclear Energy, 12*(3), 247–253. https://doi.org/10.1016/0149-1970(83)90008-2.

Lewis, D. A., & Martin, W. J. (1989). Information management: State of the art in the United Kingdom. *Aslib Proceedings, 41*(7/8), 225–250. https://doi.org/10.1108/eb051144.

Li, X., Qian, Z., You, I., & Lu, S. (2014). Towards cost efficient mobile service and information management in ubiquitous environment with cloud resource scheduling. *International Journal of Information Management, 34*(3), 319–328. https://doi.org/10.1016/j.ijinfomgt.2013.11.007.

Liu, J., Tian, J., Kong, X., Lee, I., & Xia, F. (2019). Two decades of information systems: A bibliometric review. *Scientometrics, 118*(2), 617–643. https://doi.org/10.1007/s11192-018-2974-5.

Lochbaum, K. E., & Streeter, L. A. (1989). Comparing and combining the effectiveness of latent semantic indexing and the ordinary vector space model for information retrieval. *Information Processing & Management, 25*(6), 665–676. https://doi.org/10.1016/0306-4573(89)90100-3.

López-Robles, J. R., Otegi-Olaso, J. R., Porto Gómez, I., & Cobo, M. J. (2019). 30 years of intelligence models in management and business: A bibliometric review. *International Journal of Information Management, 48*, 22–38. https://doi.org/10.1016/j.ijinfomgt.2019.01.013.

Lucas, C., Nielsen, R. A., Roberts, M. E., Stewart, B. M., Storer, A., & Tingley, D. (2015). Computer-assisted text analysis for comparative politics. *Political Analysis, 23*(2), 254–277. https://doi.org/10.1093/pan/mpu019.

Malhotra, N. K., Kim, S. S., & Patil, A. (2006). Common method variance in IS research: A comparison of alternative approaches and a reanalysis of past research. *Management Science, 52*(12), 1865–1883. https://doi.org/10.1287/mnsc.1060.0597.

Mao, H., Liu, S., Zhang, J., & Deng, Z. (2016). Information technology resource, knowledge management capability, and competitive advantage: The moderating role of resource commitment. *International Journal of Information Management, 36*(6), 1062–1074. https://doi.org/10.1016/j.ijinfomgt.2016.07.001.

Martínez-López, F. J., Merigó, J. M., Gázquez-Abad, J. C., & Ruiz-Real, J. L. (2020). Industrial marketing management: Bibliometric overview since its foundation. *Industrial Marketing Management, 84*, 19–38. https://doi.org/10.1016/j.indmarman.2019.07.014.

Martínez-López, F. J., Merigó, J. M., Valenzuela-Fernández, L., & Nicolás, C. (2018). Fifty years of the European Journal of Marketing: A bibliometric analysis. *European Journal of Marketing, 52*(1/2), 439–468. https://doi.org/10.1108/EJM-11-2017-0853.

Martorell Cunill, O., Socias Salvá, A., Otero Gonzalez, L., & Mulet-Forteza, C. (2019). Thirty-fifth anniversary of the International Journal of Hospitality Management: A bibliometric overview. *International Journal of Hospitality Management, 78*, 89–101. https://doi.org/10.1016/j.ijhm.2018.10.013.

McClure, C. R. (1981). Planning for library effectiveness: The role of information resources management. *Journal of Library Administration, 1*(3), 3–16. https://doi.org/10.1300/J111V01N03_02.

McClure, C. R., & Hill, L. L. (1982). Special considerations for corporate library planning: Moving toward information resources management. *Journal of Library Administration, 2*(2–4), 113–128. https://doi.org/10.1300/J111V02N02_09.

Mulet-Forteza, C., Martorell-Cunill, O., Merigó, J. M., Genovart-Balaguer, J., & Mauleon-Mendez, E. (2018). Twenty five years of the Journal of Travel & Tourism Marketing: A bibliometric ranking. *Journal of Travel & Tourism Marketing, 35*(9), 1201–1221. https://doi.org/10.1080/10548408.2018.1487368.

Mutch, A. (1996). Information management: A challenge for business education. *International Journal of Information Management, 16*(6), 445–455. https://doi.org/10.1016/0268-4012(96)00033-3.

Narimatsu, M. (1983). Quality information management system for purchased products. *Japan Telecommunications Review, 25*(1), 17–23. Retrieved from Scopus.

Ogiela, M. R., & Ogiela, L. (2018). Cognitive cryptography techniques for intelligent information management. *International Journal of Information Management, 40*, 21–27. https://doi.org/10.1016/j.ijinfomgt.2018.01.011.

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: An example, design considerations and applications. *Information & Management, 42*(1), 15–29. https://doi.org/10.1016/j.im.2003.11.002.

Otten, K. W. (1984). Information resources management: Management focus on the value of information and information work. *The Journal of Information and Image Management : JIIM, 17*(8), 9–14. Retrieved from Scopus.

Park, E. O., Chae, B., & Kwon, J. (2018a). Toward understanding the topical structure of hospitality literature: Applying machine learning and traditional statistics. *International Journal of Contemporary Hospitality Management, 30*(11), 3386–3411. https://doi.org/10.1108/IJCHM-11-2017-0714.

Park, E. O., Chae, B., & Kwon, J. (2018b). The structural topic model for online review analysis: Comparison between green and non-green restaurants. *Journal of Hospitality and Tourism Technology, 11*(1), 1–17. https://doi.org/10.1108/JHTT-08-2017-0075.

Parlina, A., Ramli, K., & Murfi, H. (2020). Theme mapping and bibliometrics analysis of one decade of big data research in the Scopus database. *Information, 11*(2), 1–26. https://doi.org/10.3390/info11020069.

Pasadeos, Y., Renfro, R. B., & Hanily, M. L. (1999). Influential authors and works of the public relations scholarly literature: A network of recent research. *Journal of Public Relations Research, 11*(1), 29–52. https://doi.org/10.1207/s1532754xjprr1101_02.

Paz, A. L., Merigó, J. M., Powell, P., Ramaprasad, A., & Syn, T. (2020). Twenty-five years of the Information Systems Journal: A bibliometric and ontological overview. *Information Systems Journal, 30*(3), 431–457. https://doi.org/10.1111/isj.12260.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302.

Pereira, J. V. (2009). The new supply chain's frontier: Information management. *International Journal of Information Management, 29*(5), 372–379. https://doi.org/10.1016/j.ijinfomgt.2009.02.001.

Pinho, C., Franco, M., & Mendes, L. (2018). Web portals as tools to support information management in higher education institutions: A systematic literature review. *International Journal of Information Management, 41*, 80–92. https://doi.org/10.1016/j.ijinfomgt.2018.04.002.

Pollak, M. N. (1983). Computer-aided information management systems in clinical trials. A physician's perspective. *Computer Programs in Biomedicine, 16*(3), 243–251. https://doi.org/10.1016/0010-468X(83)90086-7.

Price, D. D. S. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 292–306.

Rana, N. P., Weerakkody, V., Dwivedi, Y. K., & Piercy, N. C. (2014). Profiling existing research on social innovation in the public sector. *Information Systems Management, 31*(3), 259–273. https://doi.org/10.1080/10580530.2014.923271.

Ratner, B. (2009). The correlation coefficient: Its values range between +1/−1, or do they? Journal of Targeting. *Measurement and Analysis for Marketing, 17*(2), 139–142. https://doi.org/10.1057/jt.2009.5.

Roberts, N., & Clarke, D. (1989). Organizational information concepts and information management. *International Journal of Information Management, 9*(1), 25–34. https://doi.org/10.1016/0268-4012(89)90034-0.

Roberts, N., & Wilson, T. D. (1987). Information resource management: A question of attitudes? *International Journal of Information Management, 7*(2), 67–75. https://doi.org/10.1016/0268-4012(87)90043-0.

Roberts, M. E., Stewart, B. M., & Airoldi, E. M. (2016). A model of text for experimentation in the social sciences. *Journal of the American Statistical Association, 111*(515), 988–1003. https://doi.org/10.1080/01621459.2016.1141684.

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). STM: An r package for structural topic models. *Journal of Statistical Software, 91*(1), 1–40. https://doi.org/10.18637/jss.v091.i02.

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., … Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science, 58*(4), 1064–1082. https://doi.org/10.1111/ajps.12103.

Rockart, J. F. (1979). Chief executives define their own data needs. *Harvard Business Review, 57*(2), 81–93. Retrieved from Scopus.

Rowley, J. (1998). Towards a framework for information management. *International Journal of Information Management, 18*(5), 359–369. https://doi.org/10.1016/S0268-4012(98)00025-5.

Sambamurthy, V., Bharadwaj, A., & Grover, V. (2003). Shaping agility through digital options: Reconceptualizing the role of information technology in contemporary firms. *Management Information Systems Quarterly, 27*(2). Retrieved from https://aisel.aisnet.org/misq/vol27/iss2/6.

Saparova, D., Kibaru, F., & Bašić, J. (2013). Use of widgets as information management tools in online shared spaces. *International Journal of Information Management, 33*(2), 401–407. https://doi.org/10.1016/j.ijinfomgt.2012.12.001.

Shiau, W.-L., Dwivedi, Y. K., & Lai, H.-H. (2018). Examining the core knowledge on Facebook. *International Journal of Information Management, 43*, 52–63. https://doi.org/10.1016/j.ijinfomgt.2018.06.006. Scopus.

Silk, D. J. (1988). Towards better information management. *International Journal of Information Management, 8*(3), 167–177. https://doi.org/10.1016/0268-4012(88)90060-6.

Silva, M. M., De Gusmão, A. P. H., Poleto, T., Silva, L. C. E., & Costa, A. P. C. S. (2014). A multidimensional approach to information security risk management using FMEA and fuzzy theory. *International Journal of Information Management, 34*(6), 733–740. https://doi.org/10.1016/j.ijinfomgt.2014.07.005.

Singh, S., Dhir, S., Das, V. M., & Sharma, A. (2020). Bibliometric overview of the Technological Forecasting and Social Change journal: Analysis from 1970 to 2018. *Technological Forecasting and Social Change, 154*, Article 119963. https://doi.org/10.1016/j.techfore.2020.119963.

Soomro, Z. A., Shah, M. H., & Ahmed, J. (2016). Information security management needs more holistic approach: A literature review. *International Journal of Information Management, 36*(2), 215–225. https://doi.org/10.1016/j.ijinfomgt.2015.11.009.

Sprott, D. E., & Miyazaki, A. D. (2002). Two decades of contributions to marketing and public policy: An analysis of research published in journal of public policy & marketing. *Journal of Public Policy & Marketing*. https://doi.org/10.1509/jppm.21.1.105.17606.

Stamolampros, P., Korfiatis, N., Chalvatzis, K., & Buhalis, D. (2019). Job satisfaction and employee turnover determinants in high contact services: Insights from Employees' Online reviews. *Tourism Management, 75*, 130–147. https://doi.org/10.1016/j.tourman.2019.04.030.

Stamolampros, P., Korfiatis, Kourouthanassis, P., & Symitsi, E. (2019). Flying to quality: Cultural influences on online reviews. *Journal of Travel Research, 58*(3), 496–511.

Tranfield, D. R. (1986). Management information: Teaching at Sheffield City Polytechnic. *International Journal of Information Management, 6*(4), 257–258. https://doi.org/10.1016/0268-4012(86)90031-9.

Trauth, E. M. (1989). The evolution of information resource management. *Information & Management, 16*(5), 257–268. https://doi.org/10.1016/0378-7206(89)90003-7.

Tsay, M.-Y. (2009). Citation analysis of Ted Nelson's works and his influence on hypertext concept. *Scientometrics, 79*(3), 451–472. https://doi.org/10.1007/s11192-008-1641-7.

Tur-Porcar, A., Mas-Tur, A., Merigó, J. M., Roig-Tierno, N., & Watt, J. (2018). A bibliometric history of the journal of psychology between 1936 and 2015. *The Journal of Psychology, 152*(4), 199–225. https://doi.org/10.1080/00223980.2018.1440516.

Tvinnereim, E., & Fløttum, K. (2015). Explaining topic prevalence in answers to open-ended survey questions about climate change. *Nature Climate Change, 5*(8), 744–747. https://doi.org/10.1038/nclimate2663.

Valenzuela, L. M., Merigó, J. M., Johnston, W. J., Nicolas, C., & Jaramillo, J. F. (2017). Thirty years of the Journal of Business & Industrial Marketing: A bibliometric analysis. *Journal of Business and Industrial Marketing, 32*(1), 1–17. https://doi.org/10.1108/JBIM-04-2016-0079.

van der Aalst, W., Adriansyah, A., de Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., … Wynn, M. (2012). Process mining manifesto. In F. Daniel, K. Barkaoui, & S. Dustdar (Eds.), *Business process management workshops* (pp. 169–194). https://doi.org/10.1007/978-3-642-28108-2_19.

Van der Heijden, H. (2004). User acceptance of hedonic information systems. *Management Information Systems Quarterly, 28*(4). Retrieved from https://aisel.aisnet.org/misq/vol28/iss4/8.

Vanhala, M., Lu, C., Peltonen, J., Sundqvist, S., Nummenmaa, J., & Järvelin, K. (2020). The usage of large data sets in online consumer behaviour: A bibliometric and computational text-mining–Driven analysis of previous research. *Journal of Business Research, 106*, 46–59. https://doi.org/10.1016/j.jbusres.2019.09.009.

Wang, Y., Chen, Y., & Benitez-Amado, J. (2015). How information technology influences environmental performance: Empirical evidence from China. *International Journal of Information Management, 35*(2), 160–170. https://doi.org/10.1016/j.ijinfomgt.2014.11.005.

Wiggins, R. E. (1988). A conceptual framework for information resources management. *International Journal of Information Management, 8*(1), 5–11. https://doi.org/10.1016/0268-4012(88)90003-5.

Willard, P., & Mychalyn, J. (1998). New information management work in a changing world: An Australian survey. *International Journal of Information Management, 18*(5), 315–327. https://doi.org/10.1016/S0268-4012(98)00022-X.

Wilson, T. D. (1986). Teaching information management at the University of Sheffield. *International Journal of Information Management, 6*(4), 251–252. https://doi.org/10.1016/0268-4012(86)90027-7.

Yang, S.-N., Wang, H.-W., Gan, C.-H., & Lin, Y.-B. (2013). Mobile charging information management for smart grid networks. *International Journal of Information Management, 33*(2), 245–251. https://doi.org/10.1016/j.ijinfomgt.2012.11.007.

Yang, X., Li, Y., & Kang, L. (2020). Reconciling "doing good" and "doing well" in organizations' green IT initiatives: A multi-case analysis. *International Journal of Information Management, 51*, Article 102052. https://doi.org/10.1016/j.ijinfomgt.2019.102052.

Zheng, J., Bakker, E., Knight, L., Gilhespy, H., Harland, C., & Walker, H. (2006). A strategic case for e-adoption in healthcare supply chains. *International Journal of Information Management, 26*(4), 290–301. https://doi.org/10.1016/j.ijinfomgt.2006.03.010.

Mimno, D., Wallach, H.M., Talley, E., Leenders, M., McCallum, A. (2011). Optimizing Semantic Coherence in Topic Models. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, (EMNLP 2011), pp. 262–272. Association for Computational Linguistics, Stroudsburg. URL http://dl.acm.org/citation.cfm?id=2145432.2145462.

Bischof, J., Airoldi, E. (2012). Summarizing Topical Content with Word Frequency and Exclusivity. In J Langford, J Pineau (eds.), Proceedings of the 29th International Conference on Machine Learning (ICML 2012), pp. 201–208. Omnipress, New York.