Received: 22 July 2023 | Revised: 31 March 2024 | Accepted: 1 April 2024

DOI: 10.1002/widm.1540

ADVANCED REVIEW

WIREs DATA MINING AND KNOWLEDGE DISCOVERY

WILEY

A topic modeling-based bibliometric exploration of automatic summarization research

Xieling Chen[1] | Haoran Xie[2] | Xiaohui Tao[3] | Lingling Xu[4] | Jingjing Wang[5] | Hong-Ning Dai[6] | Fu Lee Wang[4]

[1]School of Education, Guangzhou University, Guangzhou, China

[2]Department of Computing and Decision Sciences, Lingnan University, Hong Kong, China

[3]School of Mathematics, Physics, and Computing, University of Southern Queensland, Toowoomba, Queensland, Australia

[4]School of Science and Technology, Hong Kong Metropolitan University, Hong Kong, China

[5]School of Computer Science, Hangzhou Dianzi University, Hangzhou, China

[6]Department of Computer Science, Hong Kong Baptist University, Hong Kong, China

Correspondence
Haoran Xie, Department of Computing and Decision Sciences, Lingnan University, Hong Kong, China.
Email: hrxie@ln.edu.hk

Funding information
National Natural Science Foundation of China, Grant/Award Number: 62307010; Lam Woo Research Fund of Lingnan University, Hong Kong, Grant/Award Number: LWP20019; Faculty Research Grants of Lingnan University, Hong Kong, Grant/Award Numbers: DB23A3, DR23B2

Edited by: Witold Pedrycz, Editor-in-Chief

Abstract

The surge in text data has driven extensive research into developing diverse automatic summarization approaches to effectively handle vast textual information. There are several reviews on this topic, yet no large-scale analysis based on quantitative approaches has been conducted. To provide a comprehensive overview of the field, this study conducted a bibliometric analysis of 3108 papers published from 2010 to 2022, focusing on automatic summarization research regarding topics and trends, top sources, countries/regions, institutions, researchers, and scientific collaborations. We have identified the following trends. First, the number of papers has experienced 65% growth, with the majority being published in computer science conferences. Second, Asian countries and institutions, notably China and India, actively engage in this field and demonstrate a strong inclination toward inter-regional international collaboration, contributing to more than 24% and 20% of the output, respectively. Third, researchers show a high level of interest in multihead and attention mechanisms, graph-based semantic analysis, and topic modeling and clustering techniques, with each topic having a prevalence of over 10%. Finally, scholars have been increasingly interested in self-supervised and zero/few-shot learning, multihead and attention mechanisms, and temporal analysis and event detection. This study is valuable when it comes to enhancing scholars'

Abbreviations: ACP, average citations per paper; BERT, bidirectional encoder representation from transformers; C/Y, the number of yearly citations; CNNs, convolutional neural networks; FREX, frequent and exclusive terms; GRUs, gated recurrent units; H-index, Hirsch index; LSTM, long short-term memory; MK, Mann–Kendall; NLP, natural language processing; NLTK, natural language toolkit; PRISMA, preferred reporting items for systematic reviews and meta-analyses; PTLMs, pre-trained language models; RNNs, recurrent neural networks; RQs, research questions; SNA, social network analysis; STM, structural topic model; TC, the number of total citations; TF-IDF, term frequency-inverse document frequency.

and practitioners' understanding of the current hotspots and future directions in automatic summarization.

This article is categorized under:
Algorithmic Development > Text Mining

KEYWORDS

automatic summarization, text mining, topic modeling, trend analysis

1 | INTRODUCTION

1.1 | Automatic summarization

In today's information age, a key technique is automatic text summarizing, which uses computers to create condensed versions of documents by highlighting key information from source texts (Banerjee et al., 2023). The basic objective of text summarizing is to offer a portion of the source texts that is devoid of repetition and contains the most important information (Jangra et al., 2023). Text summarization aims to decrease data while assisting users to more quickly and accurately identify and process pertinent information. The study of text summary dates back to the middle of the 20th century when Luhn (1958) first developed statistical methods such as word frequency graphs. A number of international evaluation initiatives, including the Document Understanding Conferences, the Translingual Information Detection Extraction and Summarization program in the United States, and the Text Summarization Challenge in Japan have sparked research in this area. These developments include the growth of the internet and large text databases.

It has long been believed that one must first understand the input material to generate a summary. This suggests that to determine the text's basic substance, an explicit semantic representation is required (Saggion & Poibeau, 2013). Consequently, text summarization has been acknowledged as a useful instrument for testing the understanding skills of automated systems. The complexity of the task, however, caused interest in this method of text summarization to decline, and text understanding has since emerged as a separate field of study.

Text summary has been approached from many different angles. The number of sources they employ is one method to classify them. With the information typically centered around the same subject, single-document summarization produces a summary from a single source document (Mohamed & Oussalah, 2019). Contrarily, multidocument summarization incorporates data from numerous sources that address the same subject (Christian et al., 2016; Pontes et al., 2020). SUPERT, a tool first introduced by Gao et al. (2020), assesses summary quality based on comparison with pseudo-reference summaries consisting of key phrases from the source documents. Developed by Lu et al. (2020), multi-XScience, a sizable collection of multidocument summaries of scientific papers, presents the difficult task of formulating the related work part of an article according to the abstract and references. Fuzzy models were used in a statistical feature-based multidocument summarization method developed by Patel et al. (2019) to deal with the erroneous and ambiguous feature weight. Zhang, Lu, et al. (2024) presented a summarization framework for multiple documents that follows a coarse-to-fine approach, incorporating relationships of various scales into a pipeline where extraction precedes summarization. Singh et al. (2024) developed a deep learning approach for multidocument summarization using long short-term memory (LSTM) with an improved dingo optimizer. Wahab et al. (2024) introduced a novel approach for extractive multidocument automatic text summarization based on differential evolution for multiobjective optimization and the weighted sum method.

The study of text summarizing saw a renaissance of interest in the 2000s, with a focus on creating summaries in real-time or updating ones that already exist when new information becomes available (Widyassari et al., 2022). With techniques such as online video highlighting, which uses group sparse coding to learn a dictionary from a given video and generates a summary by combining non-sparse segments (Zhao & Xing, 2014), real-time summarization techniques are utilized in various media formats, involving online videos. CatchLive is a system that uses user interaction data and stream material to produce real-time summaries of ongoing live streams, giving viewers an overview of the stream and summaries of highlight moments in a comprehensible style (Yang et al., 2022). The IncreSTS method offers an at-a-glance graphical interface for quick overview summaries and incrementally updates clustering findings with incoming comments in real-time (Liu et al., 2015). To improve document representation learning, MARES, a unique multitask learning approach for web-scale real-time event summarization, uses supervised deep learning and reinforcement learning (Yang et al., 2019).

It is essential to quickly evaluate text collections given the exponential rise of textual data online. Users would be unable to access the vast amount of information available without summaries (Saggion & Poibeau, 2013). Despite advancements in natural language processing (NLP), producing high-quality summaries of text still requires accurate text analysis, including lexical and semantic analysis. An effective summary must be accessible, succinct, non-redundant, pertinent, thorough, and contain key information (Swetha & Kumar, 2023; Zheng et al., 2010).

1.2 | Reviews of automatic summarization research

Literature reviews have been recognized as a crucial and efficient method for assessing issues and remedies within particular research fields (e.g., Abulaish et al., 2024; Ahmad et al., 2024; Yip et al., 2024). It is crucial to provide concise literature summaries since automatic summarization research is becoming increasingly significant. Widyassari et al. (2022) provided a thorough and organized evaluation of the text summarizing research conducted from 2008 to 2019. To identify and analyze research themes, tendencies, data sets, pre-processing, features, methodologies, evaluation technologies, and difficulties, 85 journal/conference papers were chosen for the review. Additionally, El-Kassas et al. (2021) provided a thorough analysis of all facets of automated text summarization, including techniques, building blocks, data sets, evaluation techniques, and future directions. Gupta and Gupta (2019) offered a thorough analysis of abstractive summarization, covering the different types of abstractive techniques used, the benefits and drawbacks of various approaches, tools created or used by researchers, evaluation methods for judging abstractive summaries, and future research directions. Allahyari et al. (2017) examined various summarizing techniques and discussed each one's merits and drawbacks. Gambhir and Gupta (2017) provided a thorough analysis of extractive summarization techniques created in the previous 10 years, covering requirements, benefits and drawbacks, abstractive and multilingual techniques, and methods of summary evaluation. An overview of extractive text summarizing approaches was presented by Moratanch and Chitrakala (2017), with an emphasis on various techniques, populated benchmarking datasets, and problems.

It is crucial to recognize that there are still several key problems regarding automatic summarization that have not been adequately addressed in the literature. For instance, what are the primary areas of focus and research interests? How have research interests evolved? Who are the influential researchers and institutions in the field? How do institutions, researchers, and countries/regions collaborate? To answer these questions, researchers must be open to adopting novel methodologies, particularly those from the field of computer science, to tackle these challenges.

1.3 | Comparing previous reviews

This study distinguishes itself from previous reviews on automatic summarization in three key aspects.

First, unlike previous studies that used qualitative analysis methodologies, the results and insights obtained in this study are not confined to pre-defined codes or categories. This is due to the utilization of topic modeling-based bibliometric methodologies, which enable the analysis of large-scale literature data. In contrast, qualitative or systematic analysis approaches typically analyze a limited number of papers, yielding results that are constrained by individual codes. For instance, Allahyari et al. (2017) performed a systematic analysis focusing on identifying summarizing techniques and discussing their merits and drawbacks. In contrast, this study identifies a wide range of automatic summarization research-related issues beyond specific techniques, as long as they are of broad interest to researchers.

Second, this study employs a nonparametric trend test to identify topics that have experienced increasing or decreasing interest in the field of automatic summarization. This type of analysis is not available in previous reviews. By examining the historical and current research progress, technological applications, and driving forces in the field, this analysis offers a comprehensive understanding of the trends. The derived results provide valuable insights and suggestions for future directions in the field, keeping automatic summarization researchers and practitioners well-informed about critical issues that require attention when pursuing scientific or technological advancements.

Furthermore, this study contributes to existing reviews on automatic summarization research by identifying key sources and contributors and visualizing research collaborations. This information will allow researchers to identify international sources that focus on automatic summarization, identify appropriate channels for making contributions, and recognize influential contributors from whom they can learn. Moreover, the study provides a clear view of the social structure and leading researchers in automatic summarization, aiding in the identification of potential academic collaborators and supporting government agencies in formulating research policies that promote the generation of knowledge in automatic summarization.

1.4 | Research aims and questions

The broad landscape and semantically significant subjects in massive amounts of textual data can be usefully represented using bibliometrics and topic models. Researchers have used these techniques to offer objective, dependable, and economically advantageous overviews of interdisciplinary study topics (e.g., Armenia et al., 2024; Chen, Xie, et al., 2023; Dwivedi et al., 2023; Sharma et al., 2021). This study combines the benefits of topic modeling and bibliometrics with a statistical trend test and social network analysis (SNA), to achieve the following objectives: (a) identify publication trends, top studies, sources, countries/regions, institutions, and researchers; (b) analyze research topics and trends in text summarization; (c) explore the development tendencies of the identified prominent topics; (d) understand how key issues and technologies in a single topic evolves; and (e) visualize the co-authorship among countries/regions/institutions/researchers. There are five research questions (RQs):

RQ1. According to the number of papers, what are the top papers, sources, countries/regions, institutions, and researchers?

RQ2. What are the hot research topics?

RQ3. How does the popularity of the topics evolve?

RQ4. How have the key issues and methodologies in each topic changed over time?

RQ5. How do institutions, researchers, and countries/regions collaborate?

The significance and drivers for exploring the RQs are manifold, demonstrating the profound insights they offer into the domain of automatic summarization.

First, by analyzing trends in publications (RQ1), we have the ability to gain valuable perspectives into the progression and evolution of automatic summarization. The upsurge in published works not only indicates the expansion of the field's scholarly landscape but also mirrors the changing dynamics (Chen, Xie, & Hwang, 2020). Identifying key contributors and publishing channels enables researchers to comprehend the global panorama of automatic summarization research, delineating developmental paths and significant stakeholders (Dong et al., 2023). This understanding assists in recognizing influential sources and fostering efficient scholarly exchange. Additionally, by discerning regional variations in research output, stakeholders can address knowledge dissemination gaps, ensuring fair access to developments in this field (Iqbal et al., 2019; Wamba et al., 2023).

Second, employing topic modeling and keyword analysis (RQ2) unveils the underlying thematic structures that are inherent in the literature of automatic summarization (Punj et al., 2023; Roberts et al., 2014). These approaches illuminate the varied spectrum of research domains and interdisciplinary intersections within the field. Identification of trending research topics enables scholars to stay updated on contemporary developments and emerging areas of exploration (Fauzi, 2022; Gurcan et al., 2021). This awareness empowers professionals to navigate scientific and technological endeavors with informed accuracy, promoting innovation and tackling research dilemmas.

Third, the incorporation of the Mann–Kendall (MK) test in examining topic dynamics (RQ3) offers a nuanced comprehension of the development of research focuses within automatic summarization (Chen, Zou, et al., 2020). Through tracking the trajectory of topic prevalence over time, scholars identify evolving patterns and persistent areas of focus (Yang, Zhang, & Rickly, 2021). These understandings shed light on the historical and present status of the field and anticipate its future trajectories (Almazroui & Şen, 2020). Foreseeing emerging trends allows stakeholders to strategically distribute resources, promoting innovation and propelling the field forward.

Fourth, observing changes in the usage of keywords and methodologies within research topics (RQ4) elucidates the evolving landscape of research interests and approaches in automatic summarization (Chen et al., 2022). This analysis uncovers emerging patterns and progressions, offering a detailed comprehension of subdomains within the field, and steering researchers toward promising avenues of inquiry and advancement (Bogers et al., 2019). Understanding the inherent dynamics of research agendas assists in identifying research priorities and informs strategic decision-making within the field.

Finally, through an investigation into research collaborations (RQ5), the research clarifies the complex web of cooperative connections among institutions, researchers, and geographical areas (Yang et al., 2023). Visualizing these collaboration networks not only pinpoints significant contributors and impactful alliances but also encourages knowledge exchange and resource sharing (Liang & Liu, 2018). Such collaborative initiatives stimulate innovation and push the field ahead, ultimately enhancing the collective progress of automatic summarization technology.

Therefore, the primary contributions of this research to the academic community can be outlined as follows: (1) introduce the first structural topic model (STM)-driven bibliometric analysis of the research field of automatic summarization; (2) uncover the key contributors (countries/regions, institutions, and authors) to share their research insights; (3) visualize collaborations among prominent contributors (countries/regions, institutions, and authors); (4) identify prevalent research topics and potential future paths; (5) enhance comprehension of the historical, current, and forthcoming academic panorama concerning automatic summarization; and (6) employ topic model-driven bibliometric methodologies for literature assessment, circumventing the constraints of manual coding or qualitative analysis techniques.

2 | DATA SELECTION

This study adhered to the three phases of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Flow Diagram (Moher et al., 2009) for identifying, selecting, and critically appraising relevant research. Additionally, a specific format was employed for collecting, analyzing, and presenting data from the studies involved in the study (Li, 2023; Pian et al., 2021). The selection of the PRISMA approach aimed to ensure rigor and minimize bias in the review process (Demir, 2021). The collected data was analyzed using topic modeling and bibliometrics, incorporating a statistical trend test and SNA. The following sections explain this process.

2.1 | Literature search

This study systematically searched the literature using five databases: Web of Science, PubMed, Scopus, ACM, and IEEE Xplore. The selection of these databases was based on their accessibility and relevance to the research theme.

The search terms used in this study were derived from previous game-based learning research (Allahyari et al., 2017; El-Kassas et al., 2021; Gambhir & Gupta, 2017; Gupta & Gupta, 2019; Moratanch & Chitrakala, 2017; Widyassari et al., 2022): ("text summarization" or "abstractive summarization" or "extractive summarization" or "document summarization" or "query-based summarization" or "generic summarization" or "automatic summarization" or "automated summarization" or "text summarisation" or "abstractive summarisation" or "extractive summarisation" or "document summarisation" or "query-based summarisation" or "generic summarisation" or "automatic summarisation" or "automated summarisation" or "text summarizer*" or "abstractive summarizer*" or "extractive summarizer*" or "document summarizer*" or "query-based summarizer*" or "generic summarizer*" or "automatic summarizer*" or "automated summarizer*" or "text summariser*" or "abstractive summariser*" or "extractive summariser*" or "document summariser*" or "query-based summariser*" or "generic summariser*" or "automatic summariser*" or "automated summariser*").

We conducted searches for the terms in titles, abstracts, and keywords, resulting in a total of 11,445 initial hits. Among these hits, there were 1785 journal papers and 2179 conference papers written in English from ACM, 4402 papers from IEEE Xplore, 2952 papers from Scopus, and 127 papers from Web of Science and PubMed databases.

2.2 | Selection criteria and study selection

The initial searches yielded a total of 11,445 papers, which were then organized in Mendeley. Within this collection, 4261 duplicate articles were automatically removed, resulting in 7184 unique papers that underwent screening based on the specified inclusion/exclusion criteria (see Table 1). To be eligible for inclusion, the studies had to meet the following criteria: (i) focus on the summarization of texts, (ii) employ text summarization methods that draw from various related NLP techniques such as text mining and text generation, and (iii) provide details about the metrics and methods used for evaluation. Studies that lacked original data (e.g., letters to the editor, commentaries, opinion pieces, theoretical articles) or were not full articles (e.g., conference abstracts, brief reports) written in English were excluded. Additionally, studies that focused on imaging and multimedia summarization or did not pertain to producing summaries using techniques such as text mining were also excluded. Finally, studies that focused solely on tools or architectures of text summarization but lacked an evaluation section were also excluded.

The initial screening of article abstracts was carried out by the first and second authors. To evaluate the consistency of selection among different coders, many reviews typically select a small random sample of abstracts to be coded by two or more coders. In this review, all retrieved abstracts were screened by two authors based on the inclusion/exclusion criteria. Whenever discrepancies arose, the third author was consulted to determine whether the abstract should be retained for full-text screening. A total of 3449 papers were excluded, resulting in 3735 papers that underwent a thorough assessment of eligibility through full-text reading. The full-text screening process was carried out by the first and second authors to finalize the inclusion of 3108 papers, following the same procedure. The inter-rater agreement reached 92%, and any discrepancies were resolved through consensus discussions. The PRISMA flow diagram in Figure 1 illustrates each stage of the search and the paper selection process.

2.3 | Data analysis for answering RQs

RQ1 was addressed by counting the number of papers and citations by year. Considering its capacity to consider non-linear associations between the year and the total papers, a polynomial modeling analysis was carried out to fit the tendencies of annual papers. The Hirsch index (H-index), average citations per paper (ACP), and other bibliometric measures were used to assess the academic performance of journals, nations/regions, academic institutions, and researchers. More specifically, the productivity and impact of actors were measured using the number of papers and citations. The total number of papers that each actor contributed was added, and the total number of citations that each paper earned was added to determine the paper count and citation count, respectively. The H-index was used to evaluate actors from both a quality and quantity perspective. The ACP of a specific actor was determined by calculating the ratio of citations to the number of papers authored.

The study used topic modeling and keyword analysis techniques to address RQ2. In addition to the specified keywords, phrases were taken from the paper titles and abstracts for keyword analysis. The researchers were able to discover frequently researched study subjects by ranking these phrases and keywords according to their frequency in the corpus of evaluated publications. We further employed topic modeling techniques to extract the semantic, intellectual frameworks, and hidden themes embedded in the dataset.

While conducting keyword analysis based on specified terms and phrases extracted from paper titles and abstracts yielded valuable insights into the content of individual research papers, it also encountered limitations (Chen, Xie, et al., 2023). Notably, numerous articles lacked keywords altogether, while some incorporated terms selected from predetermined lists provided by journals. Consequently, the keywords presented in an article may not always encapsulate

TABLE 1 Inclusion and exclusion criteria.

Inclusive criteria	I1	Studies that detail the metrics and methods used for evaluation
	I2	Text summarization methods draw from various related NLP techniques such as text mining and text generation
	I3	Studies that focus on summarization of texts
Exclusive criteria	E1	Techniques such as text mining that were not used to produce summaries
	E2	Studies focusing on tools or architectures of text summarization but lack an evaluation section
	E3	Imaging and multimedia summarization
	E4	Non-original research papers (e.g., editorials and opinion papers)
	E5	Papers not written in English

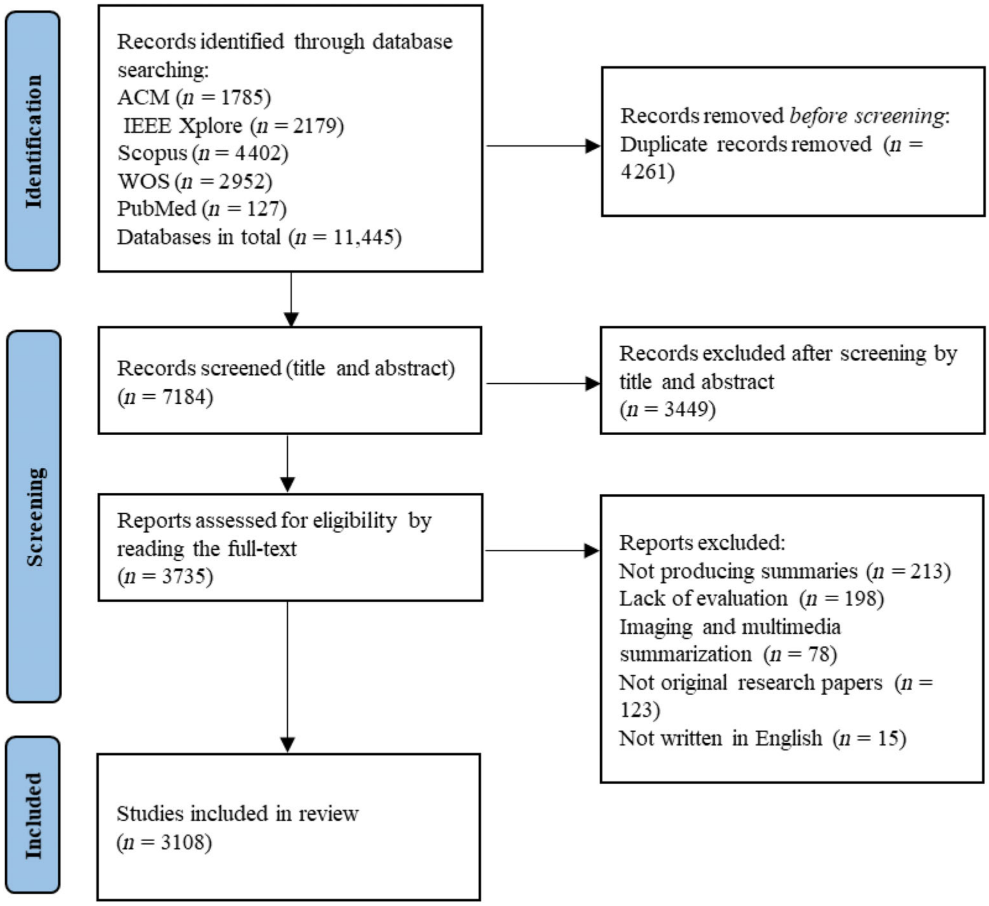


FIGURE 1 An overview of the search protocol based on the PRISMA statement.

the essence of the content optimally. Furthermore, keyword analysis represents a surface-level examination of the material (Donthu et al., 2021). While proficient at identifying specific terms and phrases within paper titles and abstracts, it may fail to capture the broader thematic context and interconnectedness between various topics. Consequently, this method might result in an incomplete comprehension of the research landscape, as it overlooks the underlying structure and trends within the evaluated publication corpus. Moreover, relying solely on keyword analysis may not adequately accommodate the variability in terminology across different research papers. Authors' choices of keywords and key phrases can vary, potentially introducing inconsistencies and biases into the analysis. Additionally, this approach may neglect emerging or less prominent topics that are not explicitly represented by the selected keywords.

To address these constraints and achieve a more thorough understanding of the prominent research themes in automatic summarization, we proceeded to conduct topic modeling analysis using an STM approach. This allowed us to delve deeper into the underlying framework of the research domain, providing a more nuanced and insightful perspective beyond the confines of individual keywords or key phrases. Scholars widely agree that, in the realm of topic discovery, topic models provide enhanced flexibility and efficacy for performing content analyses when contrasted with analyses reliant on individual terms, phrases, or keywords (Kuhn, 2018).

Figure 2 illustrates a schematic representation of the STM. Unfilled nodes symbolize latent variables, while filled nodes denote observed variables. The rectangular shapes indicate replication: $n \in \{1, 2, ..., N\}$ pertains to terms encompassing a document; $k \in \{1, 2, ..., K\}$ denotes each of the K topics; and $d \in \{1, 2, ..., \mathcal{D}\}$ signifies the document number. The primary objective of STM is to determine θ and β, which respectively signify document-topic and topic-word distributions, grounded on the terms W.

In the STM framework, θ_d signifies the hidden topic proportions for each document, while $\beta_{d,k,v}$ signifies the topic-term distributions, $z_{d,n}$ represents the fundamental assignment of topics for each term, and $w_{d,n}$ indicates the chosen term from $v \in \{1, 2, ..., V\}$. STM operates on the assumption that the creation of document d entails two stages:

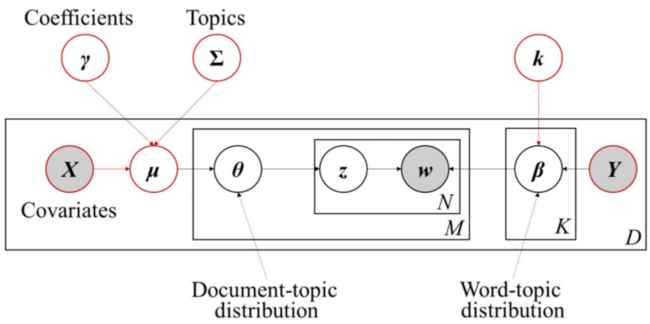


FIGURE 2 The STM diagram.

Stage 1. Randomly choose a distribution over θ_d for d.

Stage 2. For w_n in d:

a. Randomly select $z_{d,n}$ from θ_d in Step 1.
b. Randomly select a term w_n from the corresponding distribution over $\beta_{d,k,v}$, where $k = z_{d,n}$.

In this study, the STM analysis comprised three sequential steps. Initially, terms were gathered from the titles, abstracts, and keywords using the Natural Language Toolkit (NLTK) (Loper & Bird, 2002) and subsequently pre-processed to remove numbers, punctuation, and stop-words. Following this, employing term frequency-inverse document frequency (TF-IDF) technology, words deemed unimportant with a threshold of 0.05 were filtered out.

Following this, the study selected models by taking into account metrics related to exclusivity and semantic coherence. Semantic coherence, closely linked to pointwise mutual information (Lau et al., 2014; Mimno et al., 2011), was a measure employed. This metric achieves its maximum value when the most probable words within a specific topic frequently co-occur. Let $\mathcal{D}(v_i,v_j)$ represent the count of occurrences where words v_i and v_j appear together within a document. The semantic coherence for topic k is defined by Roberts et al. (2014) as Equation (1), where M signifies the top M most probable words within topic k. Each model calculates an aggregate coherence score by assessing the coherence of each topic individually and then averaging these scores.

$$C_k = \sum_{i=2}^{M} \sum_{j=1}^{i-1} \log\left(\frac{D(v_i,v_j)+1}{D(v_j)}\right). \tag{1}$$

Exclusivity quantifies the extent to which the predominant terms within a topic are distinctive to that particular topic and not widely prevalent across others. The frequent and exclusive terms (FREX) metric (Bischof & Airoldi, 2012) evaluate exclusivity by considering word frequency. FREX is defined as the weighted harmonic mean of a word's rank regarding exclusivity and frequency, as depicted in Equation (2). Here, ECDF denotes the empirical cumulative distribution function, ω represents the weight (typically set to 0.7 to prioritize exclusivity), $k \in K$ signifies the kth topic, v denotes the word under evaluation, and β refers to the topic word distribution for that specific topic. The cumulative distribution function of a real-valued random variable X, evaluated at x, denotes the probability of X assuming a value less than or equal to x. Conversely, the ECDF represents the probability distribution derived from the sampled dataset rather than the entire population.

$$\mathrm{FREX}_{k,v} = \left(\frac{\omega}{\mathrm{ECDF}\left(\beta_{k,v}/\sum_{j=1}^{K}\beta_{j,v}\right)} + \frac{1-\omega}{\mathrm{ECDF}\left(\beta_{k,v}\right)}\right)^{-1}. \tag{2}$$

In this research, the coherence and exclusivity of each topic within a model were computed using the "manyTopics" function available in the R "stm" package (Roberts et al., 2014). Subsequently, these values were averaged across all topics to determine the model's overall score. Models demonstrating high levels of exclusivity and semantic coherence were generally preferred. Out of the models with topic numbers ranging from 5 to 30, the study chose three candidate models with higher performance in terms of semantic coherence and exclusivity for manual comparison (see Figure 3).

Finally, the topic-document and term-topic percentage matrix, which demonstrated the likelihood that a term or document is relevant to a topic, was used by the researchers to compare these three candidates on an independent basis using representative terms and papers. The 14-topic model with 14 themes was found to be the best option after comparing the candidates because it included all significant topics in automatic summarization research. Following that, the label for each topic was chosen according to typical terms and papers. The FREX metric was also used to identify highly represented terms in a topic (Bischof & Airoldi, 2012). We calculated each topic's proportion using $P_k = (\sum_d \theta_{d,k})/\mathcal{D}$ to suggest their frequencies within the corpus. In this equation, P_k denotes the proportion of the kth topic, $\theta_{d,k}$ represents the proportion of the kth topic in the dth document, and $\mathcal{D}$ is 3108, representing the total number of documents.

To address RQ3, we determined the kth topic's proportion in year t using $P_{k,t} = (\sum_{d|Y=t} \theta_{d,k})/\mathcal{D}_t$ to perform a trend analysis. Here, Y_d represents the year of publication for the dth document, while $\mathcal{D}_t$ signifies the total number of documents in year t. To investigate the developmental trend of each topic, we employed an MK test (Mann, 1945). The MK test is a non-parametric method used to analyze trends in time series data.

For time series X, the test statistic S is calculated according to Equation (3):

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sign}(x_j - x_i), \tag{3}$$

$$\operatorname{sign}(x_j - x_i) = \begin{cases} -1 & \text{if } (x_j - x_i) < 0 \\ 0 & \text{if } (x_j - x_i) = 0 \\ 1 & \text{if } (x_j - x_i) > 0 \end{cases}. \tag{4}$$

In the equation, n denotes the total number of data points, x_i and x_j denote the values at times i and j (where $j > i$), respectively, and sign($x_j - x_i$) signifies the sign function using Equation (4). The test statistic S follows a normal distribution with $E(S) = 0$ and variance $V(S) = [n(n-1)(2n+5)]/18$. Z is denoted as Equation (5). A positive/negative Z indicates a rising/falling trend. With a confidence level α, a significant trend is identified when $|Z| > Z(1 - \alpha/2)$, where $Z(1 - \alpha/2)$ denotes the corresponding critical value for $p = \alpha/2$.

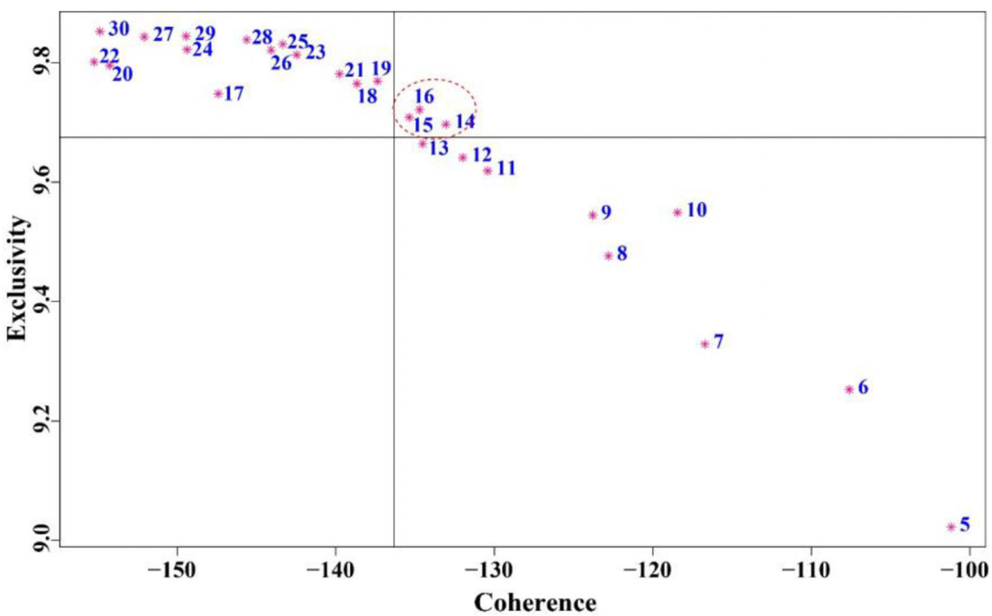


FIGURE 3 Semantic coherence and exclusivity of models with topics ranging from 5 to 30.

$$Z = \begin{cases} \frac{S-1}{\sqrt{V(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \frac{S+1}{\sqrt{V(S)}} & \text{if } S < 0 \end{cases}. \tag{5}$$

To address RQ4, we investigated the development of key issues and techniques within each topic. To determine the keywords and phrases utilized, we examined the titles and abstracts of the representative papers (having a probability of above 90%) for each topic. These were then examined using VOSviewer, with node size reflecting their frequency in the data corpus and node color representing the typical year of publication.

To respond to RQ5, we used Gephi (Bastian et al., 2009) and SNA to visually display the connections between researchers, institutions, or nations/regions by treating them as separate entities. In a cooperative network of institutions, for instance, the size of each node reflects the productivity of the institution it represents. The link's width between two nodes reveals the degree of their cooperation.

2.4 | Instruments and tools for data analysis

Table 2 describes the instruments employed for data analysis. Initially, the STM analysis was carried out using the R package "stm" (Roberts et al., 2019), utilizing the "manyTopics" function to compute both topic coherence and exclusivity. Subsequently, the MK test was conducted with the assistance of the R package "trend" (Pohlert, 2016). Third, scientific collaborations among institutions and countries/regions were visually depicted using Gephi (Bastian et al., 2009). Fourth, visualization of the annual prevalence of each topic was accomplished using the R package "ggplot" (Wickham et al., 2016). The NLTK facilitated data pre-processing tasks such as tokenization, word normalization, removal of numbers, punctuations, symbols, and stop words, as well as lemmatization for grouping different inflected forms of terms. Fifth, term filtering based on TF-IDF was executed using the R package "tm" (Feinerer et al., 2015). VOSviewer (Van Eck & Waltman, 2010) was employed as the seventh tool to visualize keywords and their evolution within each research topic. Additionally, Echarts (Li et al., 2018), a robust charting and visualization library, was utilized to display the evolution of emerging phrases, with phrases represented as nodes sized proportionally to their frequencies. Polynomial regression analysis to model the trend of annual paper numbers was conducted using the "lm" function within the R package "stats." Moreover, Python codes were developed to calculate the H-index for individual countries/regions and

TABLE 2 Instruments and tools used for data analysis.

Instruments or tools	Purposes
R package "stm"	Topic modeling analysis
manyTopics function	Coherence and exclusivity measures
R package "trend"	Mann–Kendall trend test
Gephi	Social network analysis
R package "ggplot"	Visualization of topic evolutions
NLTK	Data pre-processing (tokenization, normalization, stop word removal, lemmatization)
R package "tm"	Filtering of terms via TF-IDF
VOSviewer	Visualization of keywords and their evolutions
Python package "json"	Visualization of emerging phrases
Codes developed via Python	Calculation of the H-index
"lm" function	R package "stats"
Excel formulas such as "sum," "vlookup," "average"	Calculation of number of papers, number of citations, ACP

institutions. Excel formulas such as "sum," "vlookup," and "average" were employed to compute the number of papers, number of citations, and ACP for individual countries/regions and institutions.

3 | TREND ANALYSIS

3.1 | Publication trend

The 3108 papers were classified according to their publication year to gain insight into the progress of research on automatic summarization, as shown in Figure 4. From 2010 to 2013, we found a decrease in the number of research papers. Then, there was a slow increase until 2017. There was a sharp increase from then onwards, especially in 2018 and 2019, with a spike in 387 academic studies in 2019. In 2020, the number of automatic summarization studies showed a slight decrease, but it increased to 459 in 2021. The decrease in 2022 was caused by the incomplete coverage of automatic summarization studies because some papers were not included in indexed databases. The findings from the polynomial regression analysis indicate an exponential increase in the level of interest in the field, signifying the sustained significance and influence of research on automatic summarization in the academic domain, particularly since 2014.

3.2 | Top studies

Based on the number of total and yearly citations (TC and C/Y), Table 3 lists the top 10 research on automatic summarization among the 3092 publications (Chen et al., 2022). Nine studies (i.e., Cheng & Lapata, 2016; Dong et al., 2019; Gehrmann et al., 2018; Liu & Lapata, 2019b; Nallapati et al., 2016, 2017; Narayan et al., 2018a; Paulus et al., 2017) appear in both ranking lists. Notably, Nallapati et al.'s (2016) and Paulus et al.'s (2017) papers placed first and second, demonstrating their important contributions to research. With attentional encoder–decoder recurrent neural networks (RNNs), Nallapati et al. specifically proposed an approach for abstractive summarization that obtained cutting-edge results on two different corpora. Paulus et al. developed a neural network approach that blended reinforcement learning and conventional supervised word prediction in a novel intra-attentional framework. They tested the model using the CNN/Daily Mail and New York Times datasets, and on the CNN/Daily Mail data set, they outperformed earlier approaches with a 41.16 ROUGE-1 score.

The works on automated summarization that follow are very important. In Nallapati et al. (2017), a sequence model based on RNNs called SummaRuNNer was developed for extractive document summarization. The model allowed users

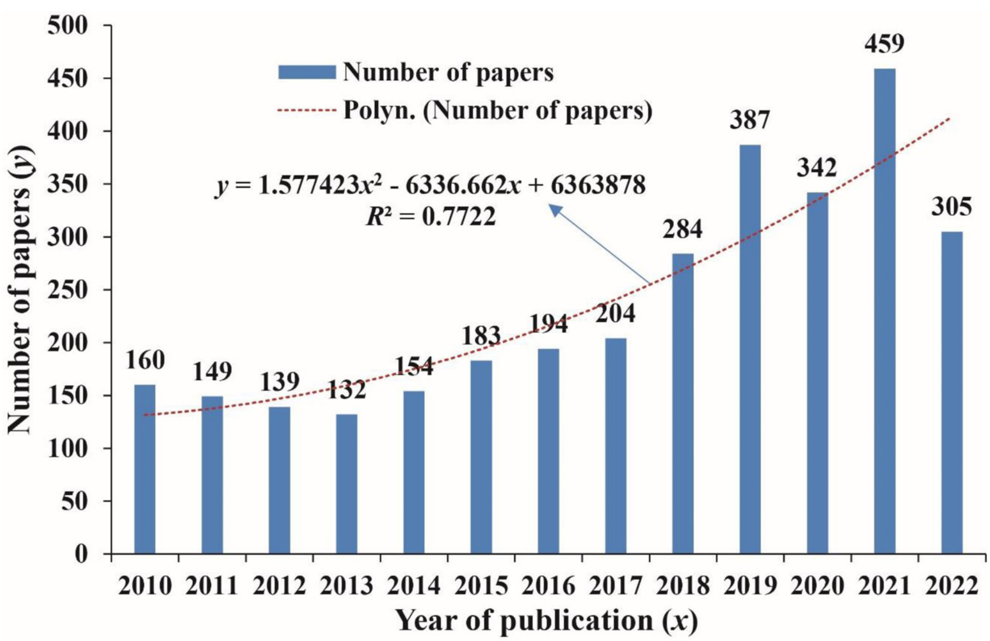


FIGURE 4 Trend analysis of the number of papers by year.

TABLE 3 Studies ranked based on the numbers of total/yearly citations.

Studies	Title	TC
Nallapati et al. (2016)	"Abstractive text summarization using sequence-to-sequence RNNs and beyond"	2017
Paulus et al. (2017)	"A deep reinforced model for abstractive summarization"	1454
Nallapati et al. (2017)	"SummaRuNNer: a recurrent neural network based sequence model for extractive summarization of documents"	1161
Liu and Lapata (2019b)	"Text summarization with pretrained encoders"	1122
Dong et al. (2019)	"Unified language model pre-training for natural language understanding and generation"	1051
Cheng and Lapata (2016)	"Neural summarization by extracting sentences and words"	833
Lin and Bilmes (2011)	"A class of submodular functions for document summarization"	829
Narayan et al. (2018a)	"Don't give me the details, just the summary! topic-aware convolutional neural networks for extreme summarization"	766
Liu et al. (2018)	"Generating Wikipedia by summarizing long sequences"	657
Gehrmann et al. (2018)	"Bottom-up abstractive summarization"	645
Studies	**Title**	**C/Y**
Paulus et al. (2017)	"A deep reinforced model for abstractive summarization"	290.80
Nallapati et al. (2016)	"Abstractive text summarization using sequence-to-sequence RNNs and beyond"	288.14
Liu and Lapata (2019b)	"Text summarization with pretrained encoders"	280.50
Dong et al. (2019)	"Unified language model pre-training for natural language understanding and generation"	262.75
Nallapati et al. (2017)	"SummaRuNNer: a recurrent neural network based sequence model for extractive summarization of documents"	193.50
Narayan et al. (2018a)	"Don't give me the details, just the summary! topic-aware convolutional neural networks for extreme summarization"	153.20
Liu et al. (2018)	"Generating Wikipedia by summarizing long sequences"	131.40
Gehrmann et al. (2018)	"Bottom-up abstractive summarization"	129.00
Cheng and Lapata (2016)	"Neural summarization by extracting sentences and words"	119.00
Chen and Bansal (2018)	"Fast abstractive summarization with reinforce-selected sentence rewriting"	110.00

Note: TC and C/Y refer to the number of total and yearly citations.

to visualize predictions based on factors including novelty, salience, and information content. A brand-new document-level encoder, based on bidirectional encoder representation from transformers (BERT), describes the semantics of a text and creates representations for each of its phrases. Some inter-sentence Transformer layers were stacked on top of this encoder to create the extractive model. The experimental findings of three datasets indicated the method's out-performance in both extractive and abstractive scenarios. Dong et al. (2019) pre-trained a unified pre-trained language model (PTLM) based on three separate language modeling tasks and used Transformer networks and self-attention mechanisms to control predictions.

There are other significant studies in automatic summarization. Using a hierarchical document encoder and attention-driven extractor, Cheng and Lapata (2016) created a generic framework for summarizing single documents. They obtained satisfactory performance by training their models on massive datasets without any language annotation. Extreme summarization, a brand-new single-document summary challenge, was described by Narayan et al. (2018a). It calls for an abstractive modeling strategy solely built on convolutional neural networks (CNNs). Both automatically and manually analyzed prevalent approaches failed to match their architecture's performance. English Wikipedia article creation may be approached as a multidocument summary problem, as shown by Liu et al. (2018), which used extractive summarization to discover significant information and a decoder-only architecture to generate abstractive models. They demonstrated how this methodology could produce whole Wikipedia papers as well as fluid, logical multisentence paragraphs. Finally, Gehrmann et al. (2018) suggested a method for selecting data-efficient material that limits the model to probable words via a bottom-up attention step. This method outperformed ROUGE for both the CNNDM and NYT corpus by a large margin.

3.3 | Analysis of publication sources

This study identified 1147 publication sources, and Table 4 presents the top 15. The Conference on Empirical Methods in Natural Language Processing and the Annual Meeting of the Association for Computational Linguistics were the most productive with 139 and 119 papers, respectively. These two sources also ranked as the top two based on H-index and citation count, followed by the Conference of the North American Chapter of the Association for Computational Linguistics—Human Language Technologies and Expert Systems with Applications. In terms of ACP value, the top three sources were the AAAI Conference on Artificial Intelligence (76.56), the Annual Meeting of the Association for Computational Linguistics (66.97), and the International Conference on Computational Linguistics (65.20), among the sources listed.

3.4 | Analyses of countries/regions, institutions, and researchers

The research on automatic summarization has received contributions from 93 countries/regions. The top 16 countries/regions are presented in Table 5. China, India, and the United States are at the top according to the number of papers, H-index, and the number of citations, highlighting their significance in this research area. With respect to ACP, the United Kingdom, Canada, and the United States are the top three in the list based on ACP value, with values of 57.09, 56.66, and 56.2, respectively.

TABLE 4 Top publication sources.

Publication sources	A (R)	H (R)	C (R)	ACP
Conference on Empirical Methods in Natural Language Processing	139 (1)	49 (1)	8933 (1)	64.27
Annual Meeting of the Association for Computational Linguistics	119 (2)	43 (2)	7969 (2)	66.97
Conference of the North American Chapter of the Association for Computational Linguistics – Human Language Technologies	66 (3)	30 (3)	3972 (3)	60.18
Expert Systems with Applications	54 (4)	30 (3)	2563 (7)	47.46
CEUR Workshop Proceedings	47 (5)	9 (19)	238 (44)	5.06
AAAI Conference on Artificial Intelligence	45 (6)	24 (6)	3445 (4)	76.56
International Conference on Computational Linguistics	45 (6)	28 (5)	2934 (5)	65.20
International ACM SIGIR Conference on Research and Development in Information Retrieval	45 (6)	20 (8)	1357 (10)	30.16
ACM International Conference on Information and Knowledge Management	40 (9)	15 (9)	982 (14)	24.55
IEEE Access	32 (10)	14 (12)	548 (18)	17.13
Information Processing and Management	29 (11)	22 (7)	1290 (11)	44.48
IEEE/ACM Transactions on Audio, Speech, and Language Processing	25 (12)	15 (9)	585 (17)	23.40
Knowledge-Based Systems	25 (12)	11 (13)	501 (20)	20.04
International Joint Conference on Neural Networks	24 (14)	4 (60)	48 (168)	2.00
International Conference on Language Resources and Evaluation	21 (15)	10 (17)	311 (31)	14.81

Note: R, H, ACP, A, and C refer to ranking position, H-index, average citations per paper, number of papers, and number of citations.

TABLE 5 Top countries/regions.

Countries/regions	A (R)	H (R)	C (R)	ACP
China	745 (1)	60 (2)	15,339 (2)	20.59
India	567 (2)	39 (3)	6595 (3)	11.63
USA	514 (3)	81 (1)	28,885 (1)	56.20
Japan	123 (4)	25 (8)	2435 (7)	19.80
UK	115 (5)	37 (4)	6565 (4)	57.09
Canada	93 (6)	31 (6)	5269 (5)	56.66
Spain	80 (7)	22 (10)	1490 (10)	18.63
Germany	76 (8)	23 (9)	1721 (9)	22.64
Iran	72 (9)	16 (15)	856 (17)	11.89
Hong Kong	70 (10)	32 (5)	3071 (6)	43.87
Indonesia	70 (10)	13 (19)	536 (26)	7.66
Taiwan	70 (10)	20 (11)	1472 (11)	21.03
Australia	68 (13)	19 (12)	1072 (14)	15.76
South Korea	67 (14)	13 (19)	685 (22)	10.22
France	62 (15)	17 (13)	1120 (13)	18.06
Malaysia	62 (15)	17 (13)	1017 (15)	16.40

Note: R, H, ACP, A, and C refer to ranking position, H-index, average citations per paper, number of papers, and number of citations.

TABLE 6 Top institutions.

Institutions	C/R	A (R)	H (R)	C (R)	ACP
Peking University	China	82 (1)	36 (1)	4380 (3)	53.41
Chinese Academy of Sciences	China	74 (2)	23 (3)	1577 (11)	21.31
Indian Institute of Technology	India	60 (3)	19 (6)	1202 (14)	20.03
Microsoft	USA	54 (4)	27 (2)	4203 (4)	77.83
University of Chinese Academy of Sciences	China	52 (5)	17 (7)	866 (19)	16.65
National Institutes of Technology	India	42 (6)	10 (27)	402 (56)	9.57
Beijing University of Posts and Telecommunications	China	41 (7)	10 (27)	358 (67)	8.73
Carnegie Mellon University	USA	35 (8)	20 (5)	2130 (8)	60.86
Hong Kong Polytechnic University	Hong Kong	34 (9)	22 (4)	1720 (10)	50.59
University of Technology Malaysia	Malaysia	30 (10)	15 (11)	780 (25)	26.00
Tsinghua University	China	29 (11)	11 (22)	525 (42)	18.10
University of Edinburgh	UK	28 (12)	17 (7)	4449 (2)	158.89
Shanghai Jiaotong University	China	26 (13)	11 (22)	494 (47)	19.00
Google	USA	25 (14)	16 (9)	1937 (9)	77.48
IBM	USA	25 (14)	16 (9)	4892 (1)	195.68

Note: R, H, ACP, A, and C refer to ranking position, H-index, average citations per paper, number of papers, and number of citations.

A total of 1779 institutions have contributed to research on automatic summarization, and the top 15 most productive institutions—six located in China and four in the United States—are included in Table 6. Regarding output as indicated by the number of papers, Peking University, the Chinese Academy of Sciences, and the Indian Institute of Technology were the top three. Regarding influence, as indicated by the H-index, the top included Peking University, Microsoft, and the Chinese Academy of Sciences. According to the number of citations, the top included IBM, the University of Edinburgh, and Peking University. As for ACP, the top in the list were IBM (195.68), the University of Edinburgh (158.89), and Microsoft (77.83).

A total of 5994 researchers have contributed to research on automatic summarization. Table 7 displays the top 14. Based on productivity and influence measured by the number of papers and H-index, the top included Furu Wei from Microsoft, Xiaojun Wan from Peking University, and Wenjie Li from Hong Kong Polytechnic University. Based on the citation count, the top three among the listed researchers were Furu Wei, Mirella Lapata from Edinburgh University, and Xiaojun Wan. According to ACP, the top in the list included Furu Wei (ACP value of 156.48), Mirella Lapata (124.14), and Sujian Li from Peking University (107.77).

3.5 | Analyses of frequently used and emerging phrases

Table 8 displays the 50 frequently used phrases, with "text summarization" appearing in 1023 papers (33.09%) as the top phrase. Other frequently used phrases are also included: "extractive summarization (510 papers, 17.49%)," "multidocument summarization (479, 15.49%)," "natural language processing (444, 14.36%)," "abstractive summarization (419, 13.55%)," and "automatic text summarization (395, 12.77%)." According to the results of the MK trend test, a significant increase in usage was observed for most of the top 50 phrases, which included "text summarization," "extractive summarization," "natural language processing," "abstractive summarization," "automatic text summarization," "computational linguistics," "neural network," and "deep learning."

Figure 5 represents the newly emerged phrases between 2018 and 2022, with an occurrence ranging from 8 to 59. Several significant emerging issues were discovered, such as "cnn/dailymail dataset," "pre-trained language model," "pointer generator network," "transformer model," "cnn/daily mail," "bidirectional encoder representation," "adversarial network," "coverage mechanism," "low-resource language," and "data augmentation."

3.6 | Analysis of topics and their trends

Table 9 displays the 14 topics acquired through the topic modeling, including their proportions and labels. Among these, the topics that received the most attention were discussed extensively: *multihead and attentions* (13.05%), *graph-based semantic analysis* (11.71%), *topic modeling and clustering techniques* (10.48%), *self-supervised and zero/few-shot learning* (7.58%), *opinion mining and personalization* (7.04%). Table 10 includes a list of representative studies that we have identified for each respective topic.

TABLE 7 Top researchers.

Researchers	Current institutions	A (R)	H (R)	C (R)	ACP
Wenjie Li	Hong Kong Polytechnic University	23 (1)	16 (3)	1398 (15)	60.78
Xiaojun Wan	Peking University	22 (2)	17 (1)	1407 (13)	63.95
Furu Wei	Microsoft	21 (3)	17 (1)	3286 (1)	156.48
Fei Liu	University of Central Florida	20 (4)	11 (11)	859 (26)	42.95
Naomie Salim	University of Technology Malaysia	17 (5)	12 (6)	535 (49)	31.47
Yogan Jaya Kumar	University of Technology Malaysia	16 (6)	9 (13)	397 (68)	24.81
Elena Lloret	University of Alicante	16 (6)	9 (13)	387 (70)	24.19
Rasim M. Alguliev	Azerbaijan National Academy of Sciences	15 (8)	13 (4)	832 (28)	55.47
Ramiz M. Aliguliyev	Azerbaijan National Academy of Sciences	15 (8)	13 (4)	832 (28)	55.47
Xiaoyan Cai	Northwestern Polytechnical University	15 (8)	8 (18)	308 (96)	20.53
Sriparna Saha	Indian Institute of Technology	15 (8)	7 (26)	163 (276)	10.87
Berlin Chen	National Taiwan Normal University	14 (12)	6 (40)	161 (279)	11.50
Mirella Lapata	Edinburgh University	14 (12)	12 (6)	1738 (9)	124.14
Sujian Li	Peking University	13 (15)	12 (6)	1401 (14)	107.77

Note: R, H, ACP, A, and C refer to ranking position, H-index, average citations per paper, number of papers, and number of citations.

TABLE 8 Top frequently used key phrases.

Key phrases	A	%	p	S	z	Trend
Text summarization	1023	33.09	0.0124	42	2.501	↑↑
Extractive summarization	510	16.49	0.0005	58	3.478	↑↑↑↑
Multidocument summarization	479	15.49	0.0005	−58	−3.478	↓↓↓↓
Natural language processing	444	14.36	0.0015	53	3.178	↑↑↑
Abstractive summarization	419	13.55	0.0001	64	3.844	↑↑↑↑
Automatic text summarization	395	12.77	0.0087	44	2.623	↑↑↑
Automatic summarization	300	9.70	0.0041	−48	−2.867	↓↓↓
Computational linguistics	297	9.61	0.0441	34	2.013	↑↑
Neural network	286	9.25	0.0008	56	3.356	↑↑↑↑
Document summarization	242	7.83	0.0087	−44	−2.623	↓↓↓
Deep learning	202	6.53	0.0000	73	4.423	↑↑↑↑
Extractive text summarization	198	6.40	0.0002	62	3.722	↑↑↑↑
Abstractive text summarization	154	4.98	0.0006	57	3.423	↑↑↑↑
Single document summarization	138	4.46	0.5022	12	0.671	↑
Information retrieval	128	4.14	0.0240	−38	−2.257	↓↓
Extractive summary	111	3.59	0.0441	34	2.013	↑↑
Abstractive summary	107	3.46	0.0001	66	3.966	↑↑↑↑
Human evaluation	101	3.27	0.0049	47	2.812	↑↑↑
Machine learning	100	3.23	0.3601	16	0.915	↑
Attention mechanism	94	3.04	0.0008	53	3.354	↑↑↑↑
News article	94	3.04	0.0995	28	1.647	↑
Unsupervised approach	92	2.98	0.2997	18	1.037	↑
Semantic similarity	90	2.91	0.9514	2	0.061	↑
Multiple document	88	2.85	0.4277	−14	−0.793	↓
Social medium	88	2.85	0.2001	22	1.281	↑
Single document	81	2.62	0.0995	28	1.647	↑
Summary sentence	80	2.59	0.2997	−18	−1.037	↓
Seq2seq model	79	2.55	0.0013	51	3.225	↑↑↑
Sentence extraction	77	2.49	0.0240	−38	−2.257	↓↓
Semantic analysis	74	2.39	0.0586	−32	−1.891	↓
Summarization system	72	2.33	0.0769	−30	−1.769	↓
Word embedding	68	2.20	0.0068	44	2.709	↑↑↑
Information extraction	67	2.17	0.0769	−30	−1.769	↓
Benchmark dataset	65	2.10	0.0995	28	1.647	↑
Clustering algorithm	65	2.10	0.1606	−24	−1.403	↓
Sentiment analysis	64	2.07	0.1424	25	1.467	↑
Sentence selection	63	2.04	0.0327	−36	−2.135	↓↓
Semantic information	62	2.01	0.0060	46	2.745	↑↑↑
Sentence ranking	62	2.01	0.0041	−48	−2.867	↓↓↓
Topic modeling	62	2.01	1.0000	0	0.000	↑
Genetic algorithm	60	1.94	0.2001	−22	−1.281	↓
Query-based summarization	60	1.94	0.2464	−20	−1.159	↓
Sentence scoring	59	1.91	0.6693	8	0.427	↑

TABLE 8 (Continued)

Key phrases	A	%	p	S	z	Trend
Information overload	57	1.84	0.5830	−10	−0.549	↓
Feature extraction	55	1.78	0.5022	12	0.671	↑
Salient sentence	54	1.75	0.6693	8	0.427	↑
Semantic relation	52	1.68	0.8548	4	0.183	↑
Sentence similarity	52	1.68	0.1272	−26	−1.525	↓
Reinforcement learning	51	1.65	0.0140	40	2.457	↑↑
Salient information	46	1.49	0.0240	38	2.257	↑↑

Note: A and % refer to the number of papers and proportions; an increasing (decreasing) trend is considered significant if $p < 0.05$. The symbols ↑↑, ↓↓, ↑↑↑, and ↓↓↓ indicate significantly increasing (decreasing) trends with *p*-values of <0.05, <0.01, and <0.001, respectively. *S* and *z* refer to the MK test and *z*-test statistics.

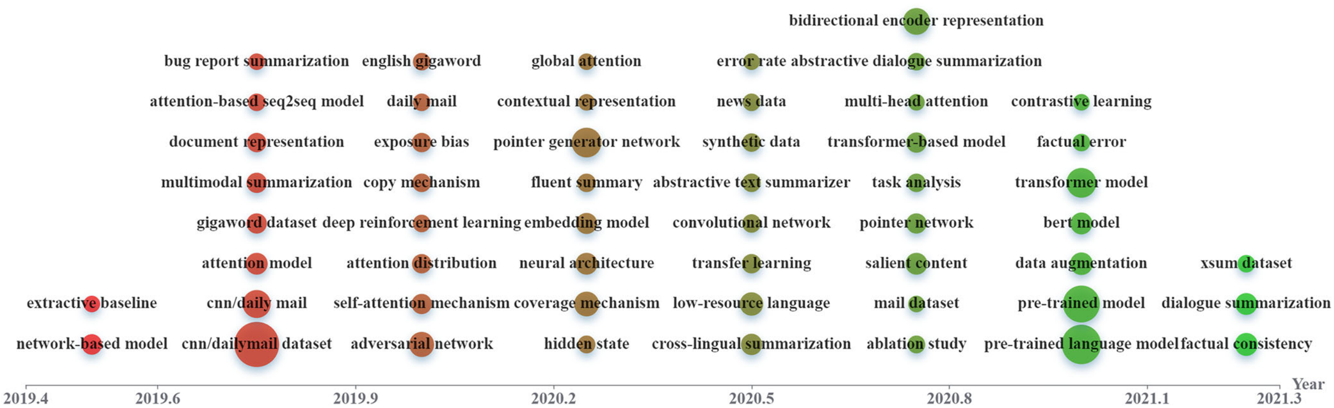


FIGURE 5 Emerging phrases between 2018 and 2022.

TABLE 9 Proportions, labels, and developmental trends for the 14 topics.

Labels	%	p	S	z	Trend
Self-supervised and zero/few-shot learning	7.58	0.000506	58	3.4775	↑↑↑↑
Multihead and attentions	13.05	0.000792	56	3.3555	↑↑↑↑
Bug reports and web documents	5.40	0.001223	−54	−3.2335	↓↓↓
Question-answering and text understanding	5.41	0.000319	−60	−3.5995	↓↓↓↓
Multimodal analysis	5.53	0.854800	−4	−0.1830	↓
Medical and clinical text summarization	6.74	0.951400	−2	−0.0610	↓
Sentence classification and compression	5.44	0.023990	−38	−2.2573	↓↓
Temporal analysis and event detection	3.62	0.017340	40	2.3793	↑↑
Multilingual and cross-lingual applications	4.59	0.200100	22	1.2812	↑
Opinion mining and personalization	7.04	0.127200	−26	−1.5252	↓
Web document summarization and disaster management	6.42	0.669300	−8	−0.4271	↓
Graph-based semantic analysis	11.71	0.032740	−36	−2.1353	↓↓
Topic modeling and clustering techniques	10.48	0.000198	−62	−3.7215	↓↓↓↓
Optimization techniques	6.97	0.023990	−38	−2.2573	↓↓

Note: "%" indicates proportion, "↑↑" (or "↓↓"), "↑↑↑" (or "↓↓↓"), and "↑↑↑↑" (or "↓↓↓↓") denote significantly increasing (or decreasing) trends with *p*-values <0.05, 0.01, and 0.001, respectively. A trend is considered not significant if $p > 0.05$. "*S*" and "*z*" represent the MK test and *z*-test statistic.

TABLE 10 Representative studies for each topic.

Topics	Studies	Title	C
Self-supervised and zero/few-shot learning	Wan and Bansal (2022)	"FACTPEGASUS: factuality-aware pre-training and fine-tuning for abstractive summarization"	14
	Fabbri et al. (2021)	"Improving zero and few-shot abstractive summarization with intermediate fine-tuning and data augmentation"	46
	Parnell et al. (2021)	"RewardsOfSum: exploring reinforcement learning rewards for summarization"	2
Multihead and attentions	Kumar et al. (2022)	"Sentic computing for aspect-based opinion summarization using multi-head attention with feature pooled pointer generator network"	1
	Yuan et al. (2020)	"Incorporating word attention with convolutional neural networks for abstractive summarization"	7
	Li et al. (2019)	"Abstractive text summarization with multi-head attention"	9
Bug reports and web documents	Mani et al. (2012)	"AUSUM: approach for unsupervised bug report summarization"	138
	Kim et al. (2019)	"A weighted PageRank-based bug report summarization method using bug report relationships"	6
	Iqbal et al. (2021)	"Big data full-text search index minimization using text summarization"	3
Question-answering and text understanding	Chan et al. (2012)	"Community answer summarization for multi-sentence question with group L1 regularization"	29
	Wang et al. (2015)	"Summarization based on task-oriented discourse parsing"	26
	Yoshida et al. (2014)	"Dependency-based discourse parser for single-document summarization"	90
Multimodal analysis	Sanabria et al. (2021)	"Hierarchical multimodal attention for deep video summarization"	8
	Sun and Tian (2022)	"Lecture video automatic summarization system based on DBNet and Kalman filtering"	7
	Li et al. (2017)	"Multi-modal summarization for asynchronous collection of text, image, audio and video"	76
Medical and clinical text summarization	Feblowitz et al. (2011)	"Summarization of clinical information: a conceptual model"	108
	Devarakonda et al. (2014)	"Problem-oriented patient record summary: an early report on a Watson application"	25
	Gulden et al. (2019)	"Extractive summarization of clinical trial descriptions"	23
Sentence classification and compression	Shams and Mercer (2015)	"Summary sentence classification using stylometry"	2
	Alias et al. (2021)	"A syntactic-based sentence validation technique for Malay text summarizer"	1
	Alias et al. (2016)	"A Malay text summarizer using pattern-growth method with sentence compression rules"	2
Temporal analysis and event detection	Li et al. (2021)	"Reinforcement learning-based dialogue guided event extraction to exploit argument relations"	22
	Koutras et al. (2015)	"Predicting audio-visual salient events based on visual, audio and text modalities for movie summarization"	27
	Lee et al. (2021)	"Event monitoring and intelligence gathering using Twitter based real-time Event summarization and pre-trained model techniques"	0

TABLE 10 (Continued)

Topics	Studies	Title	C
Multilingual and cross-lingual applications	Mrinalini et al. (2018)	"Pause-based phrase extraction and effective OOV handling for low-resource machine translation systems"	3
	Yang, Agócs, et al. (2021)	"Abstractive text summarization for Hungarian"	11
	AT et al. (2022)	"Natural language processing based cross lingual summarization"	1
Opinion mining and personalization	Porntrakoon et al. (2021)	"Text summarization for Thai food reviews using simplified sentiment analysis"	2
	Musto et al. (2019)	"Combining text summarization and aspect-based sentiment analysis of users' reviews to justify recommendations"	22
	Jiang et al. (2011)	"Capturing user reading behaviors for personalized document summarization"	1
Web document summarization and disaster management	Li and Li (2013)	"An empirical study of ontology-based multi-document summarization in disaster management"	30
	Cheng and Guo (2022)	"Automatic text summarization for public health WeChat official accounts platform based on improved TextRank"	1
	Wu et al. (2013)	"Ontology-enriched multi-document summarization in disaster management using submodular function"	12
Graph-based semantic analysis	Plaza et al. (2011)	"A semantic graph-based approach to biomedical summarization"	78
	Kumar et al. (2015)	"Graph based technique for Hindi text summarization"	22
	Han et al. (2016)	"Text summarization using FrameNet-based semantic graph model"	20
Topic modeling and clustering techniques	Cai and Li (2013)	"Ranking through clustering: an integrated approach to multi-document summarization"	55
	Wang and Zhou (2010)	"Topic-driven multi-document summarization"	20
	Cai et al. (2010)	"Simultaneous ranking and clustering of sentences: a reinforcement approach to multi-document summarization"	36
Optimization techniques	Debnath et al. (2021)	"Extractive single document summarization using multi-objective modified cat swarm optimization approach: ESDS-MCSO"	7
	Abbasi-ghalehtaki et al. (2016)	"Fuzzy evolutionary cellular learning automata model for text summarization"	63
	Ghalehtaki et al. (2014)	"A combinational method of fuzzy, particle swarm optimization and cellular learning automata for text summarization"	19

Abbreviation: C, citation count.

The MK test indicated that there was a statistically significant rise in percentage observed in three topics: *self-supervised and zero/few-shot learning*, *multihead and attentions*, and *temporal analysis and event detection*. Regarding six topics, including *bug reports and web documents*, *question-answering and text understanding*, *sentence classification and compression*, *graph-based semantic analysis*, *topic modeling and clustering techniques*, and *optimization techniques*, there was a significant increase in proportion, as demonstrated by Figure 6 which demonstrates how the 14 subjects' relative frequency in the data corpus has changed over time.

3.7 | Issues and technologies in 14 topics

Figure 7 presents the major issues/technologies and their evolutions in the 14 topics. An example of the interpretation is provided for the topic of *question-answering and text understanding*. In 2014–2015, aspect-based summarization was a commonly used approach to facilitate text understanding, with a focus on improving content-level coherence based on

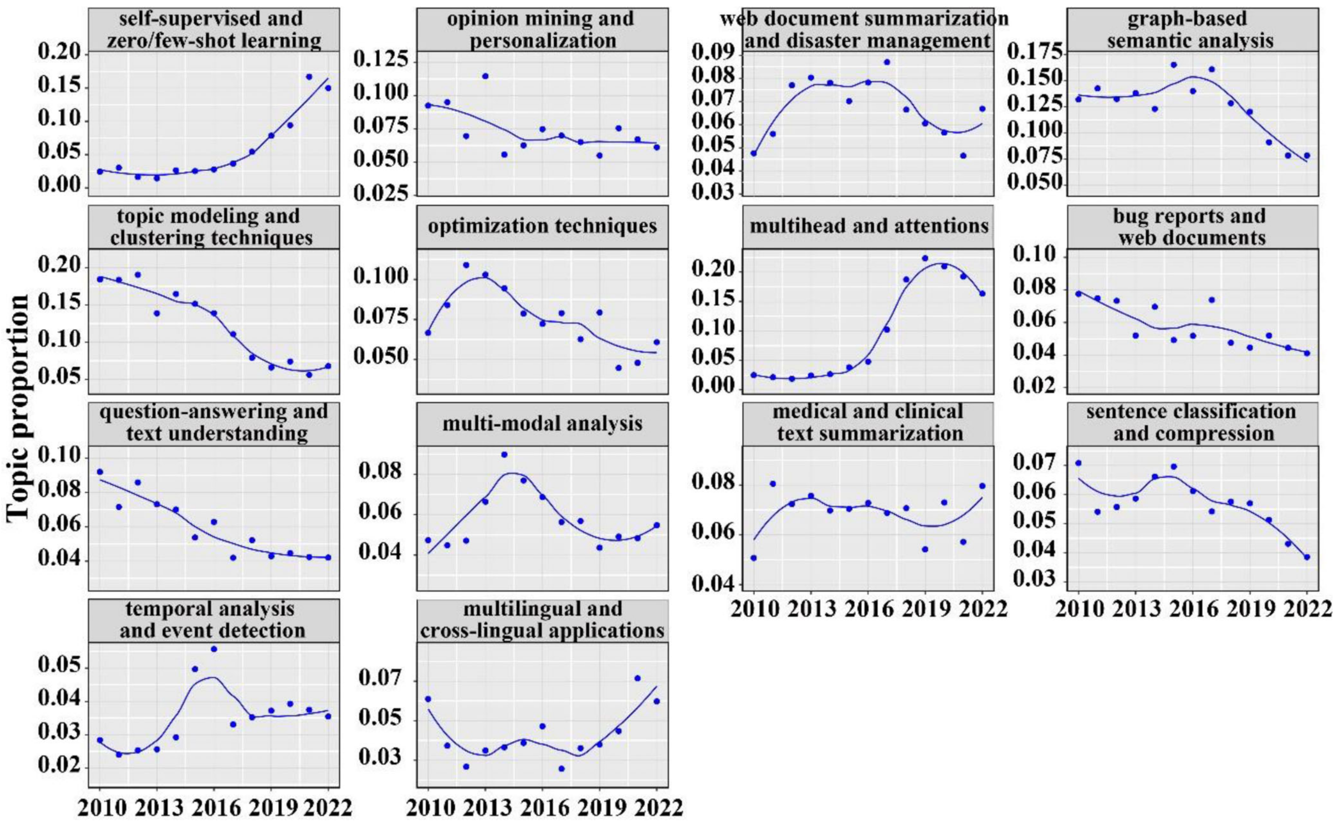


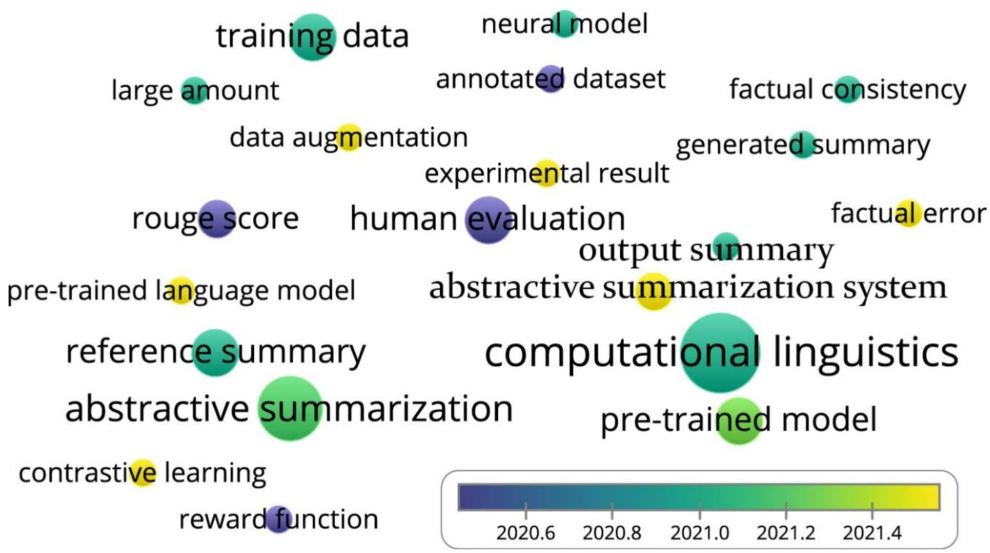
FIGURE 6 The 14 themes' annual trends.

aspect information. An approach by Zhang et al. (2013) used aspect-guided summarization, where the two prerequisite sub-tasks were identifying aspect-bearing sentences and modeling aspect-based coherence with hidden Markov models. The aspect-based models' predictions for phrase arrangement and sentence selection helped the summaries attain high coherence. Around 2015–2016, graph-based summarization methodologies gained attention for generating alternative reference texts through content summarization of top-scoring responses. Ramachandran and Foltz (2015) proposed a technique that used graph-powered cohesion technology to extract answers from among the top scorers. In recent years, deep learning advancements have stimulated research interest in automatic summarization challenges related to scientific text analysis and educational contexts. A compact meta-review that maximized information coverage, coherence, and readability, and avoided redundancy was described by Pradhan et al. (2021) using MRGen, a deep learning network-powered meta-review generating technique. The convolution layer, LSTM, bidirectional LSTM, and attention mechanism were all integrated into the proposed model MRGen to produce the final meta-review, which was based on the final decision predicted by an integrated framework.

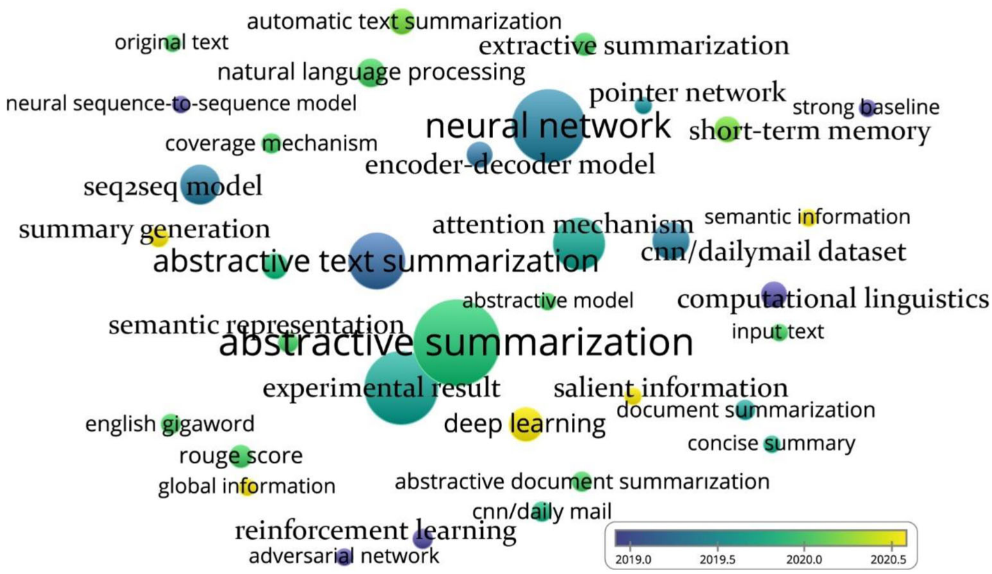
3.8 | Analysis of collaborations

Collaborations among countries/regions are illustrated in Figure 8, where the collaborative frequency ranges from 4 to 69. Nine countries/regions had collaborations with a collaborative frequency of more than 10, with the closest partners being the United States and China collaborating in 69 papers, followed by China and Hong Kong (42), and China and the United Kingdom (24). The frequency of inter-institutional collaborations ranged from 7 to 9, and four clusters were identified: (1) Japan and Viet Nam; (2) China and Canada; (3) Pakistan and Malaysia; and (4) Qatar, Brazil, the United Kingdom, the United States, Singapore, and Germany. The results showed that collaborations among countries/regions located in Asia regions are more intensive compared with other regions. For collaborations with a collaborative frequency ranging from 4 to 6, collaborative clusters formed by (1) Taiwan, China, and Saudi Arabia, as well as (2) Saudi Arabia and Pakistan are noteworthy, all of which are located in Asia.

The institutional collaborations in Figure 9 range in frequency from 4 to 40. The Chinese Academy of Sciences and the University of Chinese Academy of Sciences were the closest collaborators in 40 papers, followed by Peking



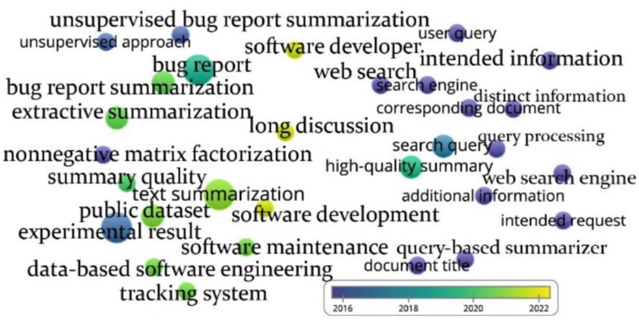
(a) Topic: self-supervised and zero/few-shot learning



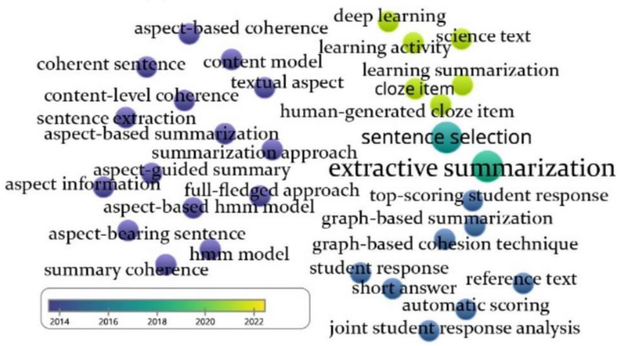
(b) Topic: multi-head and attentions

FIGURE 7 Issues/technologies in the 14 topics. (a) Topic: self-supervised and zero/few-shot learning; (b) Topic: multihead and attentions; (c) Topic: bug reports and web documents; (d) Topic: question-answering and text understanding; (e) Topic: multimodal analysis; (f) Topic: medical and clinical text summarization; (g) Topic: sentence classification and compression; (h) Topic: temporal analysis and event detection; (i) Topic: multilingual and cross-lingual applications; (j) Topic: opinion mining and personalization; (k) Topic: web document summarization and disaster management; (l) Topic: graph-based semantic analysis; (m) Topic: topic modeling and clustering techniques; (n) Topic: optimization techniques.

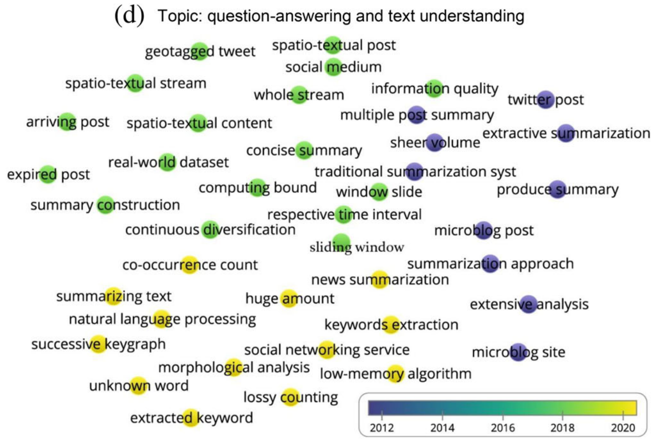
University and Hong Kong Polytechnic University (11), National Taiwan Normal University and Academia Sinica (10), and Peking University and Hong Kong Polytechnic University. Ten institutions had collaborations with a collaborative frequency of more than 8. The University of Avignon and Polytechnic School of Montreal, National Taiwan Normal University and National Taiwan University, and University of Chinese Academy of Sciences, Chinese Academy of Sciences, Aston University, and Guangzhou University were the three clusters where collaborations between institutions range in frequency from 6 to 7. The study also discovered that when compared with other collaborations, collaborations between universities from the same countries/regions were more intense. In particular, clusters formed by Guangzhou University, Peking University, and the Chinese Academy of Sciences; Guilin University of Electronic Technology and the State Key Laboratory of Mathematical Engineering and Advanced Computing; and the University of Chinese Academy of Sciences, Chinese



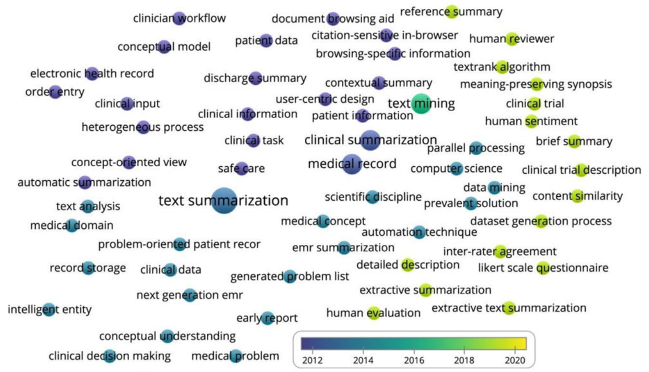
(c) Topic: bug reports and web documents



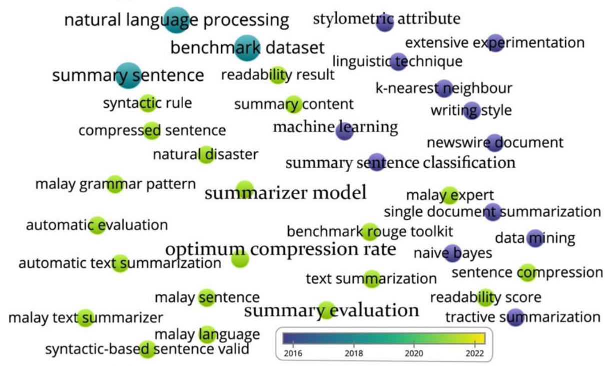
(d) Topic: question-answering and text understanding



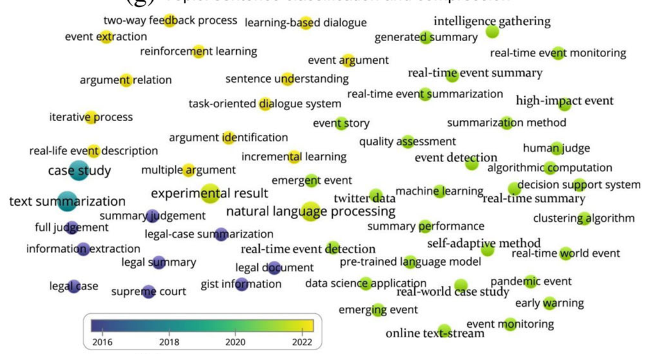
(e) Topic: multi-modal analysis



(f) Topic: medical and clinical text summarization



(g) Topic: sentence classification and compression



(h) Topic: temporal analysis and event detection

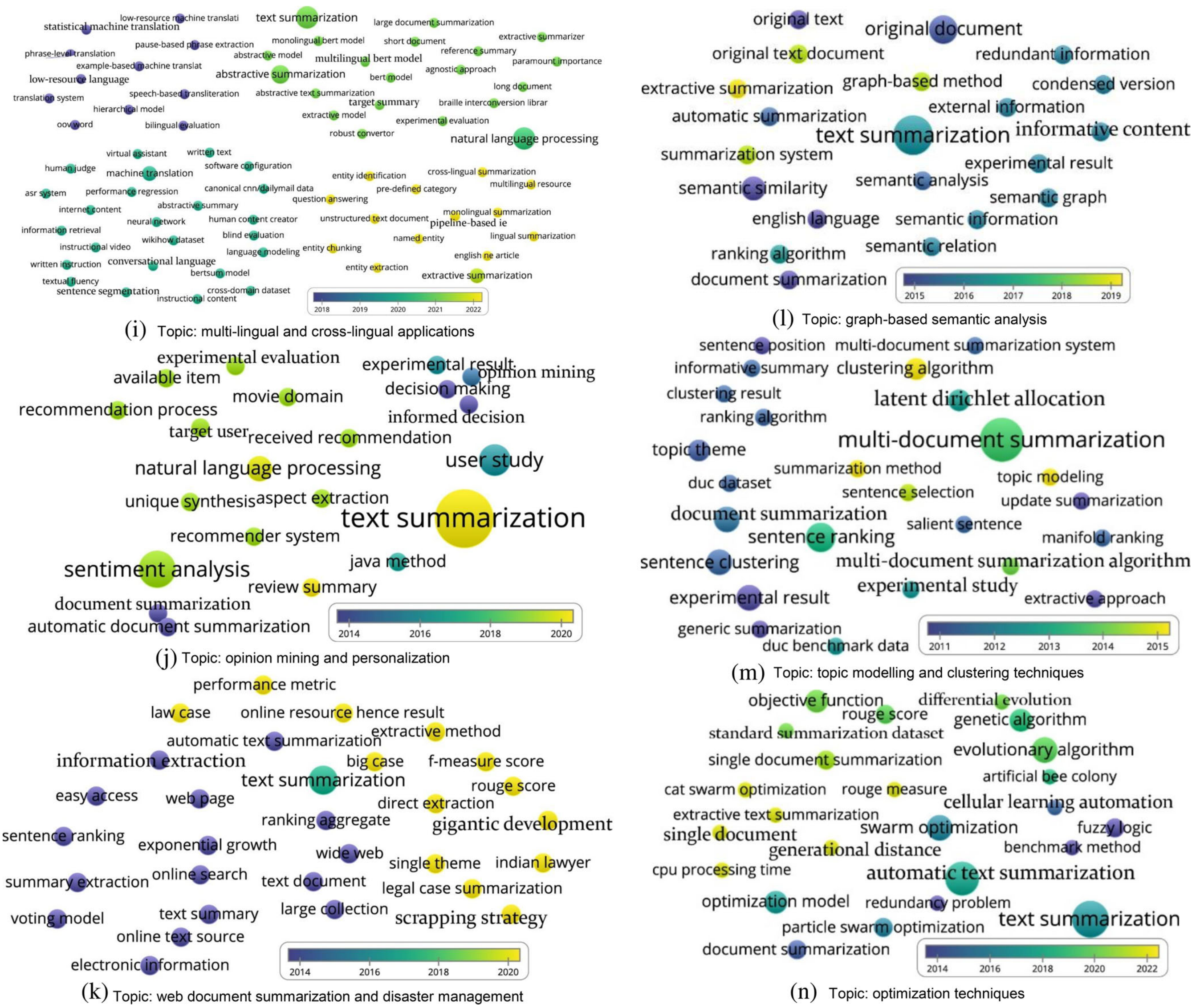
FIGURE 7 (Continued)

Academy of Sciences, Guangzhou University, Nanjing University of Posts and Telecommunications, and South China University of Technology are noteworthy for collaborations ranging in frequency from 4 to 5.

Researchers' collaborations with a collaborative frequency of 6–15 are depicted in Figure 10. Four collaborative clusters were formed by 11 researchers for collaborations with a frequency >10. These clusters include (1) Rasim M. Alguliev and Ramiz M. Aliguliyev; (2) Pushpak Bhattacharyya, Sriparna Saha, and Naveen Saini; (3) Furu Wei and Ming Zhou; as well as (4) Kuan-Yu Chen, Hsin-Min Wang, Shih-Hung Liu, and Berlin Chen. Notably, the researchers in each cluster are from the same countries/regions. This trend is further supported by collaborations with a frequency ranging from 6 to 9. For example, in the case of collaborative frequency ranging from 8 to 9, researchers from Malaysia collaborated closely in the field of automatic summarization, forming two groups, including (1) Naomie Salim and Yogan Jaya Kumar, as well as (2) Elena Baralis and Luca Cagliero.

4 | DISCUSSION

Through topic modeling and bibliometric analysis, the present study offers a comprehensive and current review of scientific research related to automatic summarization. The study covers publication trends, top studies, publication



(i) Topic: multi-lingual and cross-lingual applications

(j) Topic: opinion mining and personalization

(k) Topic: web document summarization and disaster management

(l) Topic: graph-based semantic analysis

(m) Topic: topic modelling and clustering techniques

(n) Topic: optimization techniques

FIGURE 7 (Continued)

sources, countries/regions, institutions, researchers, collaborations, prominent topics and developments, and major issues/technologies in the topics.

4.1 | In response to RQs

Regarding RQ1, the annual academic output shown in Figure 4 indicates a sustained increase in interest in automatic summarization, making it an active field in academia. Table 4 highlights that the majority of studies are published in computer science conferences, with fewer in academic journals, indicating a need for relevant international journals to organize special issues to boost outputs regarding automatic summarization. The analysis of countries/regions in Table 5 reveals that researchers from various countries, including China, India, and the United States, have shown a great interest in automatic summarization, with China contributing more than 24%. The contribution of China to automatic summarization research is also evident in the institution analysis results presented in Table 6, with Peking University, Chinese Academy of Sciences, and University of Chinese Academy of Sciences among the top five most productive institutions.

The study's response to RQ5 is supported by the network visualization presented in Figures 8–10, which indicate that countries/regions, institutions, and researchers engaging more in collaborations tend to have higher productivity

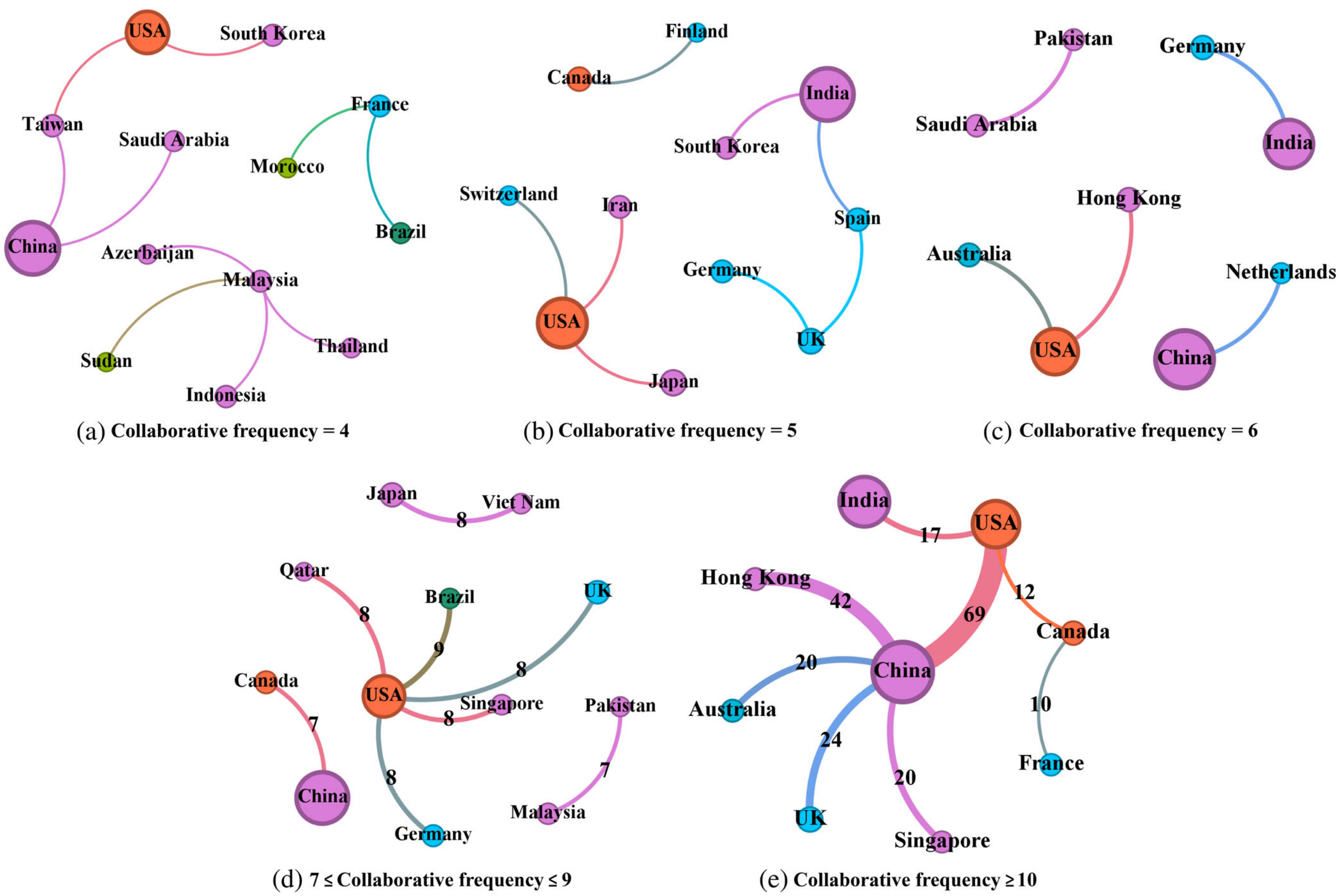


FIGURE 8 Regional collaborations range in frequency from 4 to 69.

and broader influence. Notable actors include the United States, China, and India at the regional level, Chinese Academy of Sciences and University of Chinese Academy of Sciences at the institutional level, and Furu Wei and Rasim M. Alguliev at the author level. These findings suggest international collaborations' significance in promoting this promising research field by embracing its benefits and challenges. Additionally, the study reveals that researchers from countries/regions in the same continents and institutions and researchers from the same countries/regions tend to collaborate more in automatic summarization research, similar to other scientific fields. This may be due to the convenience of resource sharing and research communication facilitated by geographical proximity. However, to promote a wider impact, the study recommends more cross-regional collaborations.

The findings presented in Tables 8 and 9 as well as Figures 5 and 6 respond to RQ2 and RQ3, indicating the prevalence of various topical groups. Three main topics, namely *multihead and attentions*, *graph-based semantic analysis*, and *topic modeling and clustering techniques*, have a proportion of over 10% each, together accounting for 35.24% of the data corpus. The results of the analysis of the frequently used and emerging phrases confirmed the popularity of these topics, with several related phrases being used frequently: "attention mechanism," "global attention," "multihead attention," "attention-based seq2seq model," "self-attention mechanism," "semantic information," "semantic analysis," "semantic similarity," "topic modeling," and "clustering algorithm." Among the three topics, *multihead and attentions* showed a significant growth trend, suggesting that it will continue to be a major focus of research. However, the other two topics, *graph-based semantic analysis* and *topic modeling and clustering techniques*, did not show any significant trend. Although these two topics received considerable attention during the study period (11.71% and 10.48%), their research interest grew slowly, indicating that their momentum may not be sustained.

Furthermore, there are five topics that make up a total of 34.75% of the data corpus, with each topic having a proportion between 6% and 10%. These topics are focused on automatic text summarization using methodologies such as self-supervised learning, zero/few-shot learning, opinion mining, and optimization for a variety of application scenarios; for example, disaster management, clinical/medical treatment and management, personalization, and web services.

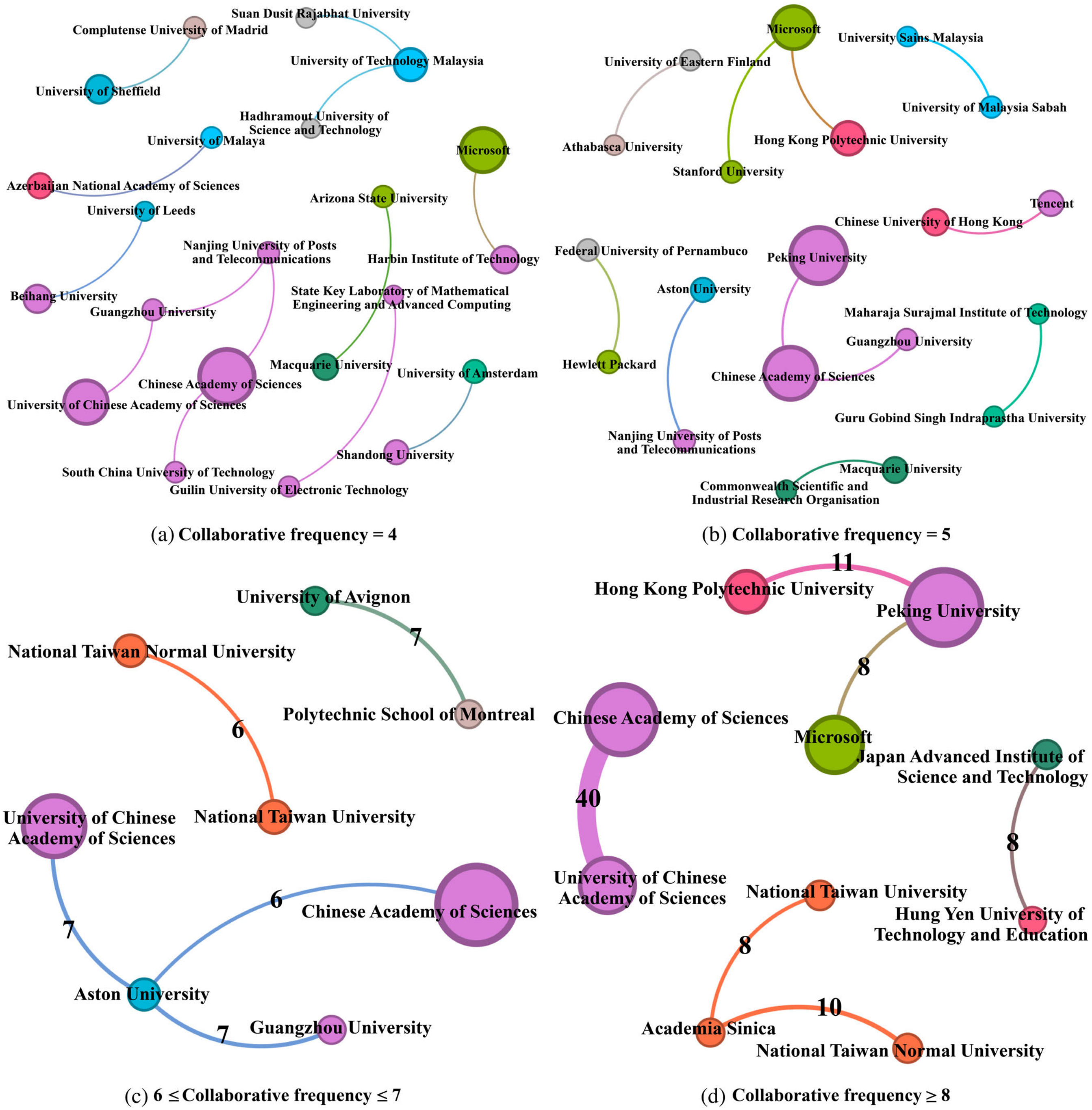


FIGURE 9 Institutions' collaborations range in frequency from 4 to 40.

Only one of them, *self-supervised and zero/few-shot learning*, has enjoyed a significantly growing tendency. This topic will likely continue to be an area of active research. However, the other four topics, such as *opinion mining and personalization*, *optimization techniques*, *medical and clinical text summarization*, and *web document summarization and disaster management*, demonstrated a declining trend in research interest, suggesting that their developmental momentum is unlikely to be sustained.

Third, the remaining six topics in the dataset have a proportion below 6%. These topics are centered around sentence classification and compression to facilitate the automatic summarization of diverse types of texts like bug reports and web documents and multilingual and cross-lingual texts to support temporal analysis and event detection, multimodal analysis, and question-answering and text understanding. The phrase analysis further supports the prevalence of these topics, with "semantic information," "cross-lingual summarization," "dialogue summarization," "abstractive

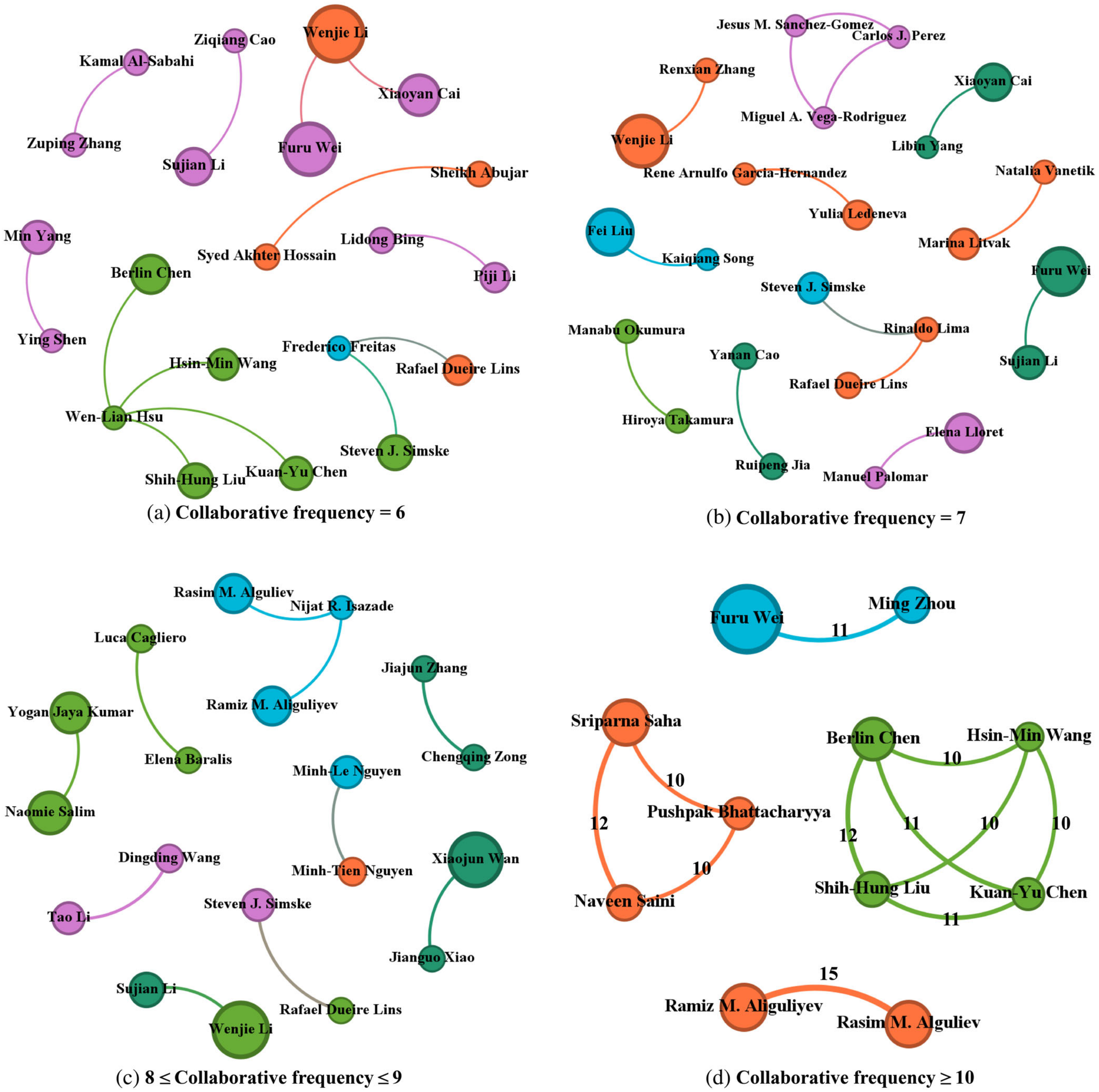


FIGURE 10 Researchers' collaborations range in frequency from 6 to 15.

dialogue summarization," "low-resource language," "cross-lingual summarization," "bug report summarization," "multimodal summarization," "query-based summarization," "clustering algorithm," and "sentence extraction." Among the six topics, only *temporal analysis and event detection* and *multilingual and cross-lingual applications* have shown growing tendencies potential for increasing interest and attention. Four topics, including *multimodal analysis*, *sentence classification and compression*, *question-answering and text understanding*, and *bug reports and web documents*, showed decreasing tendencies. The results suggest that their popularity may decrease in automatic summarization research.

The results presented in Figure 7 provide an answer to RQ2, showcasing how major issues/technologies evolved in the 14 topics. Over time, the majority of the topics witnessed the adoption of increasingly diverse technologies, with advanced machine and deep learning models being progressively utilized in various aspects of automatic summarization research. Notably, certain topics experienced a significant increase in the usage of such algorithms: *temporal analysis and event detection*, *multimodal analysis*, *question-answering and text understanding*, *multihead and attentions*, as well as *self-supervised and zero/few-shot learning*. The utilization of advanced machine learning, NLP, and deep learning algorithms (e.g., constrictive learning, PTLMs, LSTM, reinforcement learning, low-memory algorithms, and graph-based methods) has become increasingly prevalent. However, the level of diversity in terms of issues/technologies within each topic is varied. Some topics demonstrated an interest in a broad range of issues and techniques, such as *multilingual and cross-lingual applications*, *temporal analysis and event detection*, *medical and clinical text summarization*, and *multihead and attentions*. Conversely, some of the other topics showed less interest in a variety of issues and techniques, as evidenced by *self-supervised and zero/few-shot learning*, *opinion mining and personalization*, *web document summarization and disaster management*, and *graph-based semantic analysis*.

4.2 | Current status and future directions

The outcomes of topic modeling, phrase analysis, trend tests, and visualizations provide insights into the key challenges and probable directions of automatic summarization research. The frequently used and emerging phrases (Table 8 and Figure 5), combined with the four topics showing rising trends (Table 9 and Figure 6), were utilized to generate six key themes.

4.2.1 | Temporal analysis and event detection

The first theme, "temporal analysis and event detection", was formed after considering the significantly increasing importance of the topic of *temporal analysis and event detection*, which mainly focuses on incorporating temporal analysis in the summarization process and identifying important events from given texts. The ability to accurately capture and summarize information related to time and events is critical for developing effective automatic summarization systems. Temporal analysis is an essential component of automatic summarization since most of the text contains temporal information. By incorporating temporal analysis in the summarization process, we can create summaries that highlight the significant events that have occurred within a specific time frame. By identifying and extracting events, we can create summaries that are more informative and provide a better understanding of the original text. As a result, researchers have increasingly paid attention to temporal analysis and event detection as important topics in automatic summarization research (e.g., Ahmad et al., 2019; Marujo et al., 2016; Meena et al., 2023; Rajan & Jose, 2023; Sabha & Selwal, 2023). For instance, some researchers have focused on advancing automatic timeline generation by using manual mapping of timeline items to a good summary representation. The paper identified an issue with incomplete performance estimation of new timeline generation systems, which was addressed by proposed automatic solutions by McCreadie et al. (2018). The depooling methodology used by the authors showed that the risk of miss-ranking systems increased with the effectiveness of systems held out from the pool. The authors introduced two automated ground truth label expansion technologies, which reduced the number of miss-rankings by more than 50%. Additionally, some researchers focused on event extraction and detection; for example, Li et al. (2021) proposed an approach utilizing relationships between event arguments through a task-oriented dialogue system. The strategy enhanced decision-making by utilizing knowledge of previously extracted arguments and employing reinforcement and incremental learning to extract numerous arguments. The two-way feedback mechanism outperformed seven cutting-edge techniques in terms of language interpretation and event extraction.

For future work, researchers should focus on developing informative summarization systems that are able to not only extract relevant information but also understand the context and timing of that information. Moreover, summarization systems that can identify, summarize, and capture the significant and temporal events occurring over time events can have various applications in fields such as news summarization, social media analysis, and financial analysis. This can be particularly challenging in domains where events are unfolding rapidly and the relevance of information changes over time. To address this challenge, future research in automatic summarization should consider focusing on developing methods that incorporate temporal and event-based information into the summarization process. This may involve using techniques such as temporal parsing, temporal classification, and event extraction to identify important events and their associated temporal properties. Another potential avenue for future research is to explore machine learning technologies for temporal analysis and event detection in automatic summarization. This may involve developing neural network models that can learn to identify important events and their associated temporal properties from text and other data sources. Finally, it may also be beneficial to explore multimodal data sources such as texts, images, and videos for improving automatic summarization in temporal and event-based domains. This can help capture important context and timing information that may not be present in text alone.

4.2.2 | Multilingual and cross-lingual applications

The second theme "multilingual and cross-lingual applications" was formed to highlight the growing demand for summarization systems that can process and generate summaries in multiple languages, evidenced by the accelerated topic of multilingual and cross-lingual applications and the frequently used phrases such as semantic information and cross-lingual summarization. With the increasing globalization of businesses and the internet, there is a need for multilingual summarization systems that can effectively handle text in different languages. Thus, there is a need for summarization systems that can deal with multiple languages, which can be beneficial for individuals and organizations that work with multilingual documents. For example, multilingual summarization systems could be used in the translation industry to generate summaries of translated documents or to create summaries of news articles in multiple languages for global news outlets. As a result, researchers have increasingly paid attention to multilingual and cross-lingual applications as an important topic in automatic summarization research (e.g., Bhattacharjee et al., 2023; Cao et al., 2020; Shi, 2023; Takeshita et al., 2023; Xiang et al., 2024; Zheng et al., 2023). For example, Ouyang et al. (2019) provided a robust and dependable neural abstractive summarization method for cross-lingual summarization. They took advantage of machine translation and the New York Times summarizing corpus to build summarization corpora for three low-resource languages (Somali, Swahili, and Tagalog). Three abstractive summarizers were taught and tested in Arabic, a language that is not often used. The findings demonstrated that the proposed systems outperformed a standard copy-attention summarizer in terms of fluency, with comparable content selection, on automatically translated input documents.

For future work, researchers should continue to develop and test new algorithms that can effectively summarize text in multiple languages. Attention should also be paid to developing summarization systems that can handle mixed-language documents, which would be useful in situations where text in different languages is presented together. Moreover, exploring the impact of different language models and transfer learning techniques on the performance of multilingual and cross-lingual summarization systems would be an intriguing research area. Another direction could involve investigating the effect of language-specific features on summarization performance, such as the use of specific keywords or phrases that are unique to a particular language.

4.2.3 | Self-supervised and zero/few-shot learning

The third theme "self-supervised and zero/few-shot learning" was formed to highlight the potential of these techniques to enhance the quality and efficiency of automatic summarization systems, evidenced by the accelerated topic of *self-supervised and zero/few-shot learning*. Researchers have explored various approaches. One such approach is self-supervised and zero/few-shot learning techniques, which have demonstrated effectiveness in improving the performance of NLP tasks such as language modeling and machine translation (Bannur et al., 2023; Chen, Liu, et al., 2023). Self-supervised learning has the potential to improve automatic summarization by allowing the model to learn from large-scale unannotated data, which can be especially beneficial in domains where annotated data is scarce (Zhao et al., 2023). In the context of automatic summarization, self-supervised learning can be used to pre-train models on tasks like language modeling or masked language modeling, where models are trained to predict missing words in a sentence or to generate coherent text from partial input (Jiang et al., 2023). This pre-training can help the model to obtain important semantic and syntactic structures of the input texts, which can improve its ability to generate accurate and informative summaries (Zhang et al., 2023). Zero/few-shot learning is a learning paradigm that aims to develop models that generalize to new tasks with limited or no training data (Song et al., 2023). This approach can be especially useful in automatic summarization, where the amount of annotated data for a specific domain or topic may be limited. By leveraging knowledge from other related domains or topics, zero/few-shot learning can help the model generate informative summaries with limited training data (Zhang, Ladhak, et al., 2024). For instance, a model trained on news articles could be fine-tuned on scientific papers with limited training data. This approach can help the model to capture important domain-specific information and generate informative summaries.

This has led to increasing attention toward these techniques in recent years for automatic summarization tasks, as evidenced by studies (e.g., Elsahar et al., 2021; Fabbri et al., 2021; Goodwin et al., 2020; Mu et al., 2023; Wan & Bansal, 2022; Wang et al., 2019; West et al., 2019; Xie et al., 2021; Zhu et al., 2021). For example, a neural label search for summarization (NLSSum) method developed by Jia et al. (2022) learned hierarchical weights for different label sets, enabling full utilization of their information in summarization. Using both manual and automatic assessments, multilingual zero-shot summarizing studies on the MLSUM and WikiLingua data sets revealed that the NLSSum showed advanced performance. With no labeled data or fine-tuning required, Corder was a self-supervised contrastive learning system that generated vector representations of code for code retrieval tasks like code summarization (Bui et al., 2021). By applying a set of transformation operators that preserved semantics, Corder trained the source code model to distinguish between similar and different code fragments. For the tasks of code-to-code, text-to-code, and code-to-text summarization, the pre-trained models of Corder significantly outperformed baselines, giving it a practical option for eliminating labeled data in these tasks.

To summarize, by leveraging large-scale unannotated data and transferring knowledge from related domains or topics, these techniques can help to promote automatic summarization systems' robustness and adaptability. This can be particularly useful in scenarios with a limited amount of annotated data or multiple domains. Future research can focus on exploring different approaches for incorporating self-supervised and zero/few-shot learning techniques into existing summarization models. This includes exploring the impact of various pre-training tasks on summarization performance, developing evaluation metrics that can better capture the effectiveness of these techniques in real-world scenarios, and identifying the limitations and challenges of self-supervised and zero/few-shot learning for summarization.

4.2.4 | Neural network architectures and attention mechanisms

The theme “multihead and attentions” was identified as a significant aspect of improving automatic summarization models through neural network architectures and attention mechanisms, as reflected in the frequent use of phrases such as neural network, attention mechanism, and multihead attention. Recently, neural network architectures have gained popularity because of their ability to capture complicated relationships between input text and summary (Bani-Almarjeh & Kurdy, 2023; Ghadimi & Beigy, 2023; Joshi et al., 2023; Soni et al., 2023). Multihead attention, which enables a model to focus on multiple positions within the input sequence concurrently, is a promising technique for automatic summarization (Bao et al., 2023; Kumar & Solanki, 2023). By creating multiple attention heads, each learning different weights for the attention mechanism, the technique allows the model to capture more complicated relationships between the input text and summary (Jiang & Wang, 2023). This technique has been shown to improve the accuracy and informativeness of summaries.

Recently, multihead attention has gained significant attention, as evidenced by its growing usage in research (e.g., Guo et al., 2019; Kanwal & Rizzo, 2022; Li & Xu, 2023; Liu, Yang, & Cai, 2022; Wang et al., 2023). For instance, to resolve fidelity in summarizing, Liu, Yang, and Cai (2022) introduced the syntax-enriched abstractive summarizing (SEASum) paradigm that made use of graph attention networks. The architecture incorporated a GAT-powered syntactic encoder to capture explicit syntax and a PTM-driven semantic encoder to encode word sequences. Feature fusion combines encoded syntactic characteristics into summarization. Two approaches are created: parallel SEASum and cascaded SEASum. According to experimental findings on the CNN/DailyMail and Reddit-TIFU data sets, the cascaded SEASum showed better performance than traditional fidelity assessment methods. Using a multihead self-attention mechanism in the fundamental encoder–decoder approaches, Guo et al. (2019) presented the MS-Pointer Network, a deep learning method that enhanced semantic characteristics by merging input words into the encoder–decoder and giving them more weight. The approach provided a pointer network on the sequence-to-sequence with multihead attention to handle words outside of one's vocabulary and incorporated position information from the input text to improve semantic representation. Research utilizing the ROUGE measure on the CNN/DailyMail and Gigaword data sets revealed that the MS-pointer network outperformed the most recent state-of-the-art methodologies. Regarding summarizing clinical notes, Kanwal and Rizzo (2022) employed a multihead attention-based technique to extract significant words. To detect main sentences, the model connected tokens, segments, and positional embeddings of sentences. It then outputted attention ratings to extract keywords for human usage and visualization.

While multihead attention has shown some potential in automatic summarization research, there are several areas for future research. One promising path is to explore the use of multihead attention in conjunction with other techniques such as reinforcement learning or transfer learning to enhance the performance of automatic summarization approaches. Another area for future research is the development of more advanced attention mechanisms that can better capture the relationships between the input text and summary. This could involve exploring different attention functions or developing attention mechanisms that can operate on varied levels of granularity, such as word or sentence-level attention.

4.2.5 | Hybrid methods combining extractive and abstractive information

The fifth theme, titled "hybrid methods combining extractive and abstractive information," highlights the potential of hybrid summarization methods to leverage extractive and abstractive methods' advantages. Extractive and abstractive summarization are the two primary technologies for summarizing texts, each with its advantages and limitations (Mutlu & Sezer, 2023). Extractive summarization focuses on selecting critical sentences or phrases from source texts, and abstractive summarization focuses on producing new sentences to arrest the essential meaning of original texts (Ma et al., 2023). Extractive summarization is generally considered to be more reliable and easier to implement than abstractive summarization. However, abstractive summarization can produce more concise and informative summaries. To overcome the limitations of these methods, researchers have proposed to adopt hybrid approaches that combine extractive and abstractive techniques. According to Mahajani et al. (2019) and Wang et al. (2017), the aim is to leverage the advantages of both techniques, and hybrid automatic summarization systems are recommended to achieve this. For example, extractive summarization can be utilized to find the most essential sentences or phrases in original texts, which can then be used as a basis for generating an abstractive summary. Abstractive summarization can then be used to further compress the information in the summary and to generate a more concise and informative summary. Hybrid approaches can also help to overcome some of the limitations of extractive and abstractive summarization. For instance, extractive summarization may fail to detect the key ideas of the text if they are distributed across multiple sentences or paragraphs. Abstractive summarization can overcome this limitation by generating new sentences to capture the essential meanings of texts.

As a result, there is growing recognition of combining the strengths of extractive and abstractive approaches to generate more accurate and informative summaries (e.g., Alami Merrouni et al., 2023; Banerjee et al., 2023; Deng et al., 2023). For instance, Wei et al. (2019) developed a hybrid framework for single document summarization, selecting sentences with an extractive decoder and generating summaries with an abstractive decoder. BERT was used as the document encoder, with shared context representations. Experiments on the CNN/DailyMail data set demonstrated the framework's outperformance. Zaman et al. (2020) proposed a hybrid architecture by combining summarization and simplification tasks using abstractive and extractive summarization techniques. A parallel corpus was collected from simplified summaries on EurekaAlert. The model outperformed both neural text simplification and abstractive text summarization, with a new metric CSS1 showing a 38.94% and 53.40% improvement, respectively. Ghadimi and Beigy (2022) proposed HMSumm as an abstractive multidocument summarization method that combined extractive and abstractive approaches. It used a determinantal point process to handle redundancy and control input length, and a deep submodular network and BERT to generate an extractive summary. The HMSumm model used BART and T5 approaches to produce abstractive summaries, and diversity was employed to select the final summary. Experiments on DUC 2002, DUC 2004, Multi-News, and CNN/DailyMail data sets indicated that HMSumm outperformed existing methods in terms of human and ROUGE-driven evaluations.

However, hybrid approaches also face several challenges. Selecting the most critical sentences or phrases from source texts also remains as a significant challenge (Yadav et al., 2023). Extractive summarization algorithms may not always select the most relevant sentences, which influences ultimate summary quality. Another challenge is the generation of coherent and grammatically correct summaries (Wang et al., 2023). Abstractive summarization algorithms may generate sentences that are not grammatically correct or that do not convey the intended meaning. Finally, hybrid approaches require more computational resources than either extractive or abstractive summarization alone (Ramesh Kashyap et al., 2023). For future work, researchers could explore various ways to integrate extractive and abstractive methods, such as using extractive methods to identify important sentences or phrases and then rephrasing them using abstractive methods to create a more coherent and readable summary (Alami Merrouni et al., 2023; Goyal et al., 2023; Wu, 2024). Additionally, researchers could examine how different weighting schemes or scoring mechanisms for combining extractive and abstractive approaches may affect the generated summary quality. Another direction could be to explore the impact of hybrid approaches on various applications such as information retrieval and text summarization for social media or online news.

4.2.6 | State-of-the-art technologies and mechanisms

The sixth theme "state-of-the-art technologies and mechanisms" was formed to highlight the cutting-edge performance and accuracy demonstrated by prevalent technologies and mechanisms in automatic summarization, evidenced by frequently used and emerging phrases such as PTLMs, transfer learning, pointer generator networks, Transformers, attention-based seq2seq models, coverage mechanism, and copy mechanisms.

Deep neural sequence-to-sequence

Deep neural sequence-to-sequence approaches are currently frequently employed in NLP missions such as automatic text summarization, machine translation, and information retrieval (Babu & Badugu, 2023; Baykara & Güngör, 2023; Carichon et al., 2023). These models are designed to convert input sequences into desired outputs, such as translating a text sequence between two languages. The main approach for creating sequence-to-sequence approaches is known as encoder–decoder architectures, where the encoder network creates a hidden representation of inputted sequences utilized by decoder networks to create outputs (Dalmia et al., 2023). However, deep sequence-to-sequence models encounter challenges such as handling long-term dependencies, being difficult to parallelize, having low novelty, exposure-bias problems, evaluation mismatch, and lacking generalization. To tackle these challenges, attention, copy, and coverage mechanisms have been used to strengthen the basic deep neural sequence-to-sequence architectures, which result in significant improvements in the performance of automatic summarization tasks.

Recurrent, convolutional, and attention

In both extractive and abstractive summarization, RNNs and CNNs are adopted (Aliakbarpour et al., 2024; Giarelis et al., 2023; Li et al., 2024). IBM has built models for abstractive summarization based on RNNs, while Facebook has developed a model based on CNNs. The Transformer design has shown to have considerable increases in performance and ROUGE ratings. PTLMs and other models have used this attention-based architecture, which was designed exclusively based on attention processes, as an encoder and/or decoder.

The self-attention mechanism of the Transformer has overcome the issues of sequential nature and long-term dependency of RNN by modeling input tokens' similarities despite their positions (Qin et al., 2024; Rahman et al., 2024). Consequently, Transformer-powered approaches have shown reliable results on various automatic summarization tasks (Bani-Almarjeh & Kurdy, 2023; Kumar & Solanki, 2023; Searle et al., 2023) and outperformed RNN/CNN-powered sequence-to-sequence approaches with less training time. For example, a highlighting mechanism was added to the encoder in KPAT, proposed by Liu, Cao, et al. (2022), to assign higher attention weights to tokens within main phrases. Key phrases were extracted and scored, and a highlighting matrix was built to indicate their importance. Two highlighting attention structures were designed and tested on different datasets, where KPAT significantly outperformed advanced summarization baselines. A neural summarization approach was developed by Liu and Lapata (2019a) that processed multiple input documents and generated abstractive summaries. The Transformer architecture was augmented with hierarchical document encoding and attention mechanisms to represent cross-document relationships. Furthermore, the Transformer-based approach used explicit graph representations for similarity or discourse connections and might discover latent correlations between textual elements. The HIBERT algorithm proposed by Zhang et al. (2019) leveraged hierarchical BERTs and was pre-trained on unlabeled input. When compared with randomly initialized versions, it performed significantly better, achieving 1.25 ROUGE on CNN/DailyMail and 2.0 ROUGE on a New York Times dataset.

In the field of NLP, there is a growing trend of modifying the internal architecture of Transformers by combining them with recurrent units such as gated recurrent units (GRUs) and LSTM. This approach has resulted in significant improvements in both speed and performance across various automatic summarization tasks. For instance, Adelia et al. (2019) utilized bidirectional GRUs–RNNs approaches to improve text summarization in Bahasa Indonesia, which was commonly achieved using extractive methods with low inter-sentence cohesion. Similarly, Gambhir and Gupta (2022) employed convolutional Bi-GRUs to extract syntactic/semantic relationships from texts.

Beyond deep learning: Reinforcement learning, transfer learning, and PTLMs

To address RNN-powered approaches' limitations in NLP, such as difficulty in handling long-range dependencies and inability to work in parallel resulting in low-quality summaries, the reinforcement learning and transfer learning approaches are combined (Uc-Cetina et al., 2023; Xiong et al., 2024). Low novelty, exposure bias, assessment mismatch, and a lack of generalization are further problems that these models encounter. With the help of reinforcement learning, abstractive automatic summarization results have been improved in terms of ROUGE scores, factual consistency, readability, coherency, and syntax (Atri et al., 2023; Frisoni et al., 2023). Rich semantic and contextual features offered by PTLMs, which pre-train Transformers on large corpora and transfer the knowledge to downstream tasks like abstractive summarization, enhance the quality of the resulting summaries (Zhang, Lu, et al., 2024).

In several automatic summarization research fields, reinforcement learning and PTLMs have achieved noteworthy results. For instance, Li et al. (2021) used reinforcement and incremental learning to obtain arguments using an iterative, multiturn procedure to capture argument connection. A new goal function was proposed by Gao et al. (2018) that permitted using active, preference, and reinforcement learning strategies to decrease sample complexity. Extractive summarization was treated by Narayan et al. (2018b) as a sentence ranking problem, and it introduced a unique training approach by using reinforcement learning to globally maximize the ROUGE evaluation measure. To tackle low-resource abstractive summarization, Chen and Shuai (2021) suggested using pre-trained approaches and diverse corpora, where pre-trained models enhanced primary summarization ability and diverse corpora improved generalization. Experiments conducted on various corpora validated the approach. A transfer learning-based method was developed by Givchi et al. (2022) to generate abstractive summaries from extractive summaries. It addressed the challenge of interpreting essential concepts and generating new paraphrased sentences not identical to the main text.

Technologies for long summarization

Automatic text summarization research's focus has traditionally been generating short summaries from news article datasets, but there is now a growing interest in summarizing long documents (Joshi et al., 2023; Zheng et al., 2023). However, summarizing long documents is more complicated and requires more efficient approaches due to the complex hardware requirements. Applying models designed for short summaries to long documents can lead to incoherent and trivial summaries.

Long texts can be divided into numerous encoders or sub-models using reinforcement learning techniques to provide coherent summaries (Aliakbarpour et al., 2024; Alomari et al., 2022; Yao et al., 2018), but these models become increasingly complex and are not parallelizable for longer documents. To address this issue, researchers suggest using Transformers that can handle long-term dependencies more effectively, because of the ease with which gradient flow occurs due to their logarithmic or constant route length (Bani-Almarjeh & Kurdy, 2023). Transformer-powered approaches are designed to address long documents effectively by modeling commonalities between words that are independent of their placements utilizing the self-attention mechanism (Li et al., 2023). For instance, Beltagy et al. (2020) introduced Longformer, a linearly scalable attention mechanism that could handle lengthy documents with thousands of tokens. Longformer's attention mechanism replaced the standard self-attention by combining windowed and task-motivated global attention. Longformer was evaluated on character-level language modeling, pre-trained, and fine-tuned on several downstream tasks, outperforming RoBERTa on long document tasks and setting new benchmarks on WikiHop and TriviaQA. Furthermore, the Longformer–encoder–decoder had been introduced for summarizing lengthy documents. Grail et al. (2021) proposed a hierarchical propagation layer that distributed information between transformer windows by dividing the input into multiple blocks processed independently by scaled dot-attentions and combined between layers. The approach achieved satisfactory performance for long scientific papers and news articles' extractive summarization and comparable outcomes for shorter documents, surpassing language-model-based summarizers.

Pointer-generator/copy and coverage mechanisms

By integrating extractive and abstractive goals, attention-based sequence-to-sequence approaches with copy mechanisms address the shortcomings of unfamiliar words and erroneous details in text summaries (Ma et al., 2023; Shi et al., 2021). Researchers have improved the copy mechanism by using a directed graph and the self-attention layer of the Transformer to examine the degree of centrality of single source words, allowing for exact control of the copy process. Although useful, these models frequently produce summaries with repeated sentences, which lessens their novelty levels. The coverage method, which tracks the created information to prevent duplication, was designed to overcome this problem. The creation of repetitive sentences, which lowers the novelty levels in the summaries, is a typical problem despite the benefits of adopting copy processes in attentional sequence-to-sequence models. The coverage mechanism has been introduced as a solution to this issue (Aliakbarpour et al., 2022; Babu & Badugu, 2023; Li et al., 2024). This prevents the repetition of information by keeping track of the created content in the summary. For providing relevant and succinct solutions to non-factoid inquiries in question-driven summarization, as an illustration, gated selective pointer generation networks with multiview coverage methods were proposed by Deng et al. (2020).

4.3 | Insights and implications

4.3.1 | Insights

In terms of annual output, this study reveals a 65% growth in the number of automatic summarization papers, rising from 160 in 2010 to 459 in 2021. This finding aligns with the research conducted by Widyassari et al. (2022), who identified two papers in 2008 and 18 papers in 2018. Notably, the numbers reported in Widyassari et al.'s study were considerably lower compared with those in this study. This discrepancy could be attributed to their utilization of only ScienceDirect, IEEE, and ACM databases, while our study incorporates a broader range of databases for data search. Nonetheless, the increase in automatic summarization literature indicated by both Widyassari et al.'s research and our study suggest that this field is experiencing growth, which is characterized by rapid development and significant potential.

Regarding publication sources, this study has identified that automatic summarization research is particularly well-received in computer science conferences, while fewer papers are published in journals. This finding contrasts with the observations made by Widyassari et al., who found that Expert Systems with Applications was the most prolific journal in terms of publishing automatic summarization studies. The disparity primarily arises from the fact that Widyassari et al. included 80% of eligible journal papers and only 20% of conference papers related to automatic summarization in their review, whereas this study encompasses all eligible papers. Consequently, our results suggest that the research community in this field places significant value on conference publications as the primary avenue for sharing and disseminating their work.

Regarding collaborators, our results indicate that Asian countries and institutions, notably China and India, are among the top two contributors to automatic summarization. They actively engage in the field, with their contributions surpassing 24% and 20% of the total output, respectively, ranking even higher than the third contributor, the United States. However, in a previous review conducted by Mishra et al. (2014), they identified the United States as the dominant actor, accounting for 47% of their analyzed data, while countries/regions from Asia only contributed 6%. This difference could be attributed to variations in data coverage. Specifically, Mishra et al. focused on publications before 2014, whereas our study covers papers from 2010 to 2022. These findings suggest that the substantial increase in the number of automatic summarization papers is primarily due to significant contributions made by Asian countries/regions and institutions.

Regarding scientific collaborations, the network visualization results indicate that countries/regions and institutions displaying a strong inclination toward international collaboration demonstrate higher productivity and have a wider impact. Notably, China and India serve as prime examples in this regard. The collaborations appear to be more prevalent among countries/regions and institutions located in the same regions, while cross-regional collaborations are less pronounced.

Regarding research topics and trends, this study contributes significantly by providing results that go beyond pre-defined codes. For instance, the findings are more specific and detailed and can capture the latest trends in automatic summarization research. This distinguishes our study from previous reviews on automatic summarization and related topics.

First, it is worth noting that results from previous reviews employing qualitative analysis methodologies often tend to be limited to individual pre-defined codes. For example, Widyassari et al.'s systematic analysis of 85 papers focused on identifying techniques and methods utilized in automatic text summarization. However, by employing STM-based bibliometric methodologies capable of analyzing extensive literature data, this study identifies a broader range of issues related to automatic summarization research. These issues are not confined to specific methods commonly used to address automatic summarization problems. Instead, if certain issues garner widespread attention among automatic summarization researchers, they are identified in this study. Examples of such identified issues include problem-related aspects (e.g., optimization, personalization, question-answering, and event detection), method-related aspects (e.g., multihead and attention mechanisms, semantic analysis), and application domain-related aspects (e.g., medical and clinical applications, disaster management, multilingual and cross-lingual capabilities).

Second, in comparison to previous studies that have employed qualitative analysis approaches, this study offers valuable insights into specific and nuanced research themes. These insights are achieved through a collaborative interpretation and refinement of results obtained from topic modeling and phrase analysis, combined with a meticulous examination of papers that focus on the identified topics and phrases. For instance, one of the prominent research themes highlighted in this study is the exploration of "hybrid methods combining extractive and abstractive information" for text summarization. This theme emphasizes the potential of hybrid summarization methods that harness the strengths of both extractive and abstractive approaches. The identification of this theme provides a more specific, detailed, and actionable perspective compared with the general issues mentioned in previous reviews. Widyassari et al. discovered that most automatic summarization studies were centered around abstractive and extractive methods. However, no information regarding the combination of these two methods was provided. In contrast, our study delves deeper into the specific research theme of hybrid methods, shedding light on the opportunities and advantages associated with integrating extractive and abstractive techniques in the context of automatic summarization.

Furthermore, in comparison to previous reviews, this study offers deeper insights into the latest trends in automatic summarization research. For instance, previous reviews (e.g., Gupta & Gupta, 2019; Mishra et al., 2014; Widyassari et al., 2022) suggested the dominance of machine learning, TF-IDF, and fuzzy approaches, while overlooking the significance of NLP, semantic analysis, topic modeling, graph-based techniques, and neural network approaches, particularly deep learning. However, this study identifies the prevalence and growing interest in utilizing these advanced technologies within the realm of automatic summarization research. Specifically, we observe the application of graph-based semantic analysis, topic modeling, clustering techniques, multihead attention mechanisms, and seq2seq models. Moreover, Gambhir and Gupta suggested developing hybrid methods that combine extractive and abstractive techniques, whereas our study reveals that the use of hybrid methods has gained popularity and is known for producing high-quality summaries. Similarly, Gupta and Gupta highlighted the future direction of cross-language and multilingual summarization. In contrast, our study identifies automatic summarization as a vital component in many multilingual and cross-lingual applications. These discrepancies can be attributed to the wider coverage of relevant and up-to-date papers in our study compared with previous reviews. Additionally, the lack of analysis regarding the developmental trends and emerging issues of specific topics in previous reviews, in contrast to our study, further contributes to these differences. Consequently, we can identify the latest trends in automatic summarization research that have seldom been mentioned in previous reviews, such as the use of automatic summarization for personalization and disaster management.

4.3.2 | Implications

Based on the findings and discussions, the implications can be summarized as follows:

First, the rapid development and potential of automatic summarization offer abundant research opportunities for new researchers and PhD students. Exploring various aspects of automatic summarization, such as novel algorithms, evaluation methods, and applications in specific domains, particularly through interdisciplinary approaches, collaboration with experts from different fields, and integration of insights from multiple domains such as NLP, machine learning, information retrieval, and cognitive science, can lead to impactful contributions in the field.

Second, international journals specializing in related disciplines (e.g., NLP, machine learning, and information retrieval) can play a crucial role in promoting automatic summarization research. They can achieve this by launching special issues to encourage researchers from diverse backgrounds to contribute their expertise and foster cross-disciplinary collaboration.

Third, researchers, especially newcomers, can look to the main contributors identified in this study as potential role models from whom they can learn and establish collaborations. Both intra-regional and cross-regional collaborations can help researchers effectively explore opportunities and address the challenges in automatic summarization.

Fourth, researchers are advised to continue focusing on topics that receive wide and increased interest, such as multihead and attention mechanisms, graph-based semantic analysis, topic modeling and clustering techniques, self-supervised and zero/few-shot learning, and temporal analysis and event detection.

Fifth, as automatic summarization becomes increasingly integrated into practical applications, collaborations between scholars and professionals from computer science, linguistics, healthcare, and other relevant domains should be strengthened. This will facilitate the development of effective real-world summarization systems tailored to specific domains or use cases, including multilingual and cross-lingual analysis, disaster management, personalization, and medical and clinical support.

Finally, researchers should leverage state-of-the-art technologies, such as self-supervised and zero/few-shot learning, multihead attention mechanisms, hybrid methods, deep neural sequence-to-sequence approaches, GRUs, reinforcement learning, transfer learning, PTLMs, pointer-generator/copy mechanisms, and coverage mechanisms. By utilizing these technologies and their combinations, researchers can facilitate the development of effective real-world applications, such as temporal event detection, personalized services, disaster management, intelligent health and medical systems, and question-answering systems.

4.4 | Limitations, reflections, and future work

The current study investigates the research topics, contributors, and collaborations of automatic summarization research using bibliometrics and topic models. There are three limitations to take into consideration. Regarding data collection and selection, only journal articles and conference papers were included. To obtain a deeper grasp of the area of automated summary research, additional research should be performed with the inclusion of diverse resources such as books and dissertations. Although they were retrieved via search string sharing across many research topics, a large number of publications were omitted from manual data evaluation since they were unrelated to automatic summarization. Context-specific queries should be taken into account while modifying search strategies in order to improve future work.

In terms of methodology, this study evaluated the performance of publication sources, nations/regions, institutions, and authors using citation-based bibliometric measures including citation count and H-index. However, due to various factors that can affect academic influences, such as the establishment, novelty, or interdisciplinary nature of a journal, caution must be exercised in interpreting these results. This was addressed by using alternate metrics such as the number of papers and ACP to assess these entities from various angles.

Regarding topic analysis, this research employed a topic modeling-driven bibliometrics approach instead of relying on systematic reviews, meta-analyses, or the involvement of domain experts in crafting the summary. While involving domain experts in crafting the summary might appear as a straightforward approach, it presents practical hurdles and may not be feasible given the breadth and scale of our investigation. Systematic reviews and meta-analyses concentrate on scrutinizing specific and pre-defined codes or categories within a limited number of articles. Given the broad spectrum and extensive volume of publications, alongside the dynamic nature of the automatic summarization field, involving domain experts in crafting the summary would necessitate experts to individually assess each paper and subsequently summarize the topics, which is a process that would be exceedingly time-consuming and labor-intensive. Furthermore, systematic reviews and meta-analyses often demand a substantial quantity of homogeneous data, which might be challenging to acquire in a swiftly evolving and multidisciplinary domain like automatic summarization. The diversity in research methodologies, terminologies, and focal points across various publications could present difficulties in conducting a comprehensive systematic review or meta-analysis. Moreover, the direct subjective evaluation by experts on a large dataset could yield divergent outcomes.

Conversely, STM provides an automated and data-centric methodology for summarizing extensive text datasets. By utilizing statistical algorithms and NLP methods, STM discerns latent topics within the dataset by analyzing patterns and co-occurrences of words. This methodology allows for the systematic exploration of recurrent themes and emerging trends in the realm of automatic summarization, eliminating the necessity for manual intervention or the subjective assessment of each paper.

By utilizing automatic topic modeling methodology, machines are adept at handling vast volumes of literature data within a short timeframe. Subsequently, experts are tasked with evaluating the estimated terms and documents exhibiting high likelihood and exclusivity to topics to perform labelling. Grounded on the objective outcomes of topic models, the summary of topics elucidates crucial and recurring themes as well as emerging trends in automatic summarization research from both quantitative and qualitative standpoints. Moreover, achieving consensus in comparing the 14 labels is relatively straightforward; however, attaining consensus becomes challenging when dealing with hundreds or even thousands of labels provided by diverse experts. Hence, employing topic models, as opposed to systematic analysis methodologies, facilitates the attainment of more effective, efficient, and objective results.

Furthermore, topic modeling empowers us to capture the multidimensional and intricate nature of the research landscape, encompassing both overarching thematic categories and specific subtopics within automatic summarization. This level of granularity in analysis would pose challenges for traditional systematic or meta-analysis methodologies, which might overlook nuanced variations in research themes and priorities. In the future, it would be compelling to undertake a systematic analysis of the papers to gain more detailed insights into the field's evolution. This endeavor would necessitate the development of techniques enabling the automatic execution of systematic analysis on a large dataset.

However, it is recognized that while STM autonomously generates topics, it still necessitates expert input to evaluate the output quality. The participation of domain experts in generating summaries based on topic modeling results is undeniably beneficial and can significantly enhance the interpretability and accuracy of the topics generated. Integrating domain experts into the process can provide additional insights, enhance the coherence of topics, and ensure alignment with the nuances of the field. Our aim was not to supplant expert involvement, but rather to streamline the initial exploration phase, thereby optimizing the utilization of experts' time and resources. Presently, no study interprets results without human intervention. In future studies, exploring the potential for proposing automatic methodologies to interpret topic modeling results would be intriguing.

In the current investigation, the experts' assessment of topic modeling outcomes has reaffirmed its capability to encompass most topics deemed significant by domain experts. However, achieving comprehensive coverage of all aspects within a field is inherently challenging due to the presence of overlapping topics, and conceptually ambiguous words may result in certain issues going undetected. Conceptually spurious words are those that may have multiple interpretations and can pose challenges under certain circumstances. This phenomenon is a common challenge encountered in topic modeling research. Additionally, achieving comprehensive coverage of all facets of every study, including those relying on direct expert summaries using a systematic analysis approach involving manual evaluation of extensive datasets, is difficult. As the topic interpretation process adhered strictly to prior research protocols, the interpretation outcomes are deemed acceptable. Nonetheless, in future research, it would be interesting to complement text mining with comprehensive systematic analysis to yield more nuanced results.

Although our present study offers a thorough examination of research topics and trends in automatic summarization, we acknowledge the importance of conducting more detailed inquiries into specific subdomains within the field. In future studies, we aim to explore opportunities for conducting more targeted investigations into particular subareas of automatic summarization. Through deeper exploration of individual topics and comparative analysis of these individual investigations, we can furnish a more nuanced comprehension of the field and the nuances of each research area, thereby contributing to the advancement of knowledge in this domain.

5 | CONCLUSION

This work used topic modeling and bibliometrics to analyze the scientific literature in order to clarify subjects and their developments in research on automatic summarization. This work explored topic dynamics through a statistical trend test, identified and displayed topic distributions across major players, and indicated research frontiers in addition to doing so. According to an analysis of the yearly output, researchers' interest in this area was growing. China, India, and the United States were among the most productive nations and areas. The Chinese Academy of Sciences, Peking University, and Indian Institute of Technology were among the most productive institutions. International collaborations facilitated quicker development and improved academic performance. Graph-based semantic analysis, topic modeling, multihead and attentions, and clustering approaches were among the topics that were widely discussed. There was a statistically significant rise in interest in the research areas of self-supervised and zero/few-shot learning, multihead and attentions, and temporal analysis and event detection. This study adds to the body of knowledge about automatic summarization. It offers significant and useful insights, especially for helping academics, decision-makers, and practitioners better understand the general setting and organization of this increasingly active field. The discovered productive actors may be used by researchers as possible partners and role models. Additionally, to further explore the advantages and get past the difficulties of automatic summarization, especially those empowered by deep neural networks, and to promote the best possible decision-making, academic cooperation can be expanded and extended.

AUTHOR CONTRIBUTIONS

Xieling Chen: Conceptualization (equal); data curation (equal); formal analysis (equal); investigation (equal); methodology (equal). **Haoran Xie:** Formal analysis (equal); funding acquisition (equal); methodology (equal); project administration (equal). **Xiaohui Tao:** Investigation (equal); methodology (equal). **Lingling Xu:** Data curation (equal); formal analysis (equal). **Jingjing Wang:** Data curation (equal); formal analysis (equal). **Hong-Ning Dai:** Formal analysis (equal). **Fu Lee Wang:** Funding acquisition (equal); investigation (equal); project administration (equal).

ACKNOWLEDGMENTS

The research has been supported by the National Natural Science Foundation of China (No. 62307010), as well as the Direct Grant (DR23B2), Lam Woo Research Fund of Lingnan University, Hong Kong (LWP20019), and the Faculty Research Grants of Lingnan University, Hong Kong (DB23A3 and DR23B2).

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

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

ORCID

Haoran Xie https://orcid.org/0000-0003-0965-3617
Xiaohui Tao https://orcid.org/0000-0002-0020-077X
Fu Lee Wang https://orcid.org/0000-0002-3976-0053

RELATED WIREs ARTICLES

On the application of machine learning in astronomy and astrophysics: A text-mining-based scientometric analysis

REFERENCES

Abbasi-ghalehtaki, R., Khotanlou, H., & Esmaeilpour, M. (2016). Fuzzy evolutionary cellular learning automata model for text summarization. *Swarm and Evolutionary Computation*, *30*, 11–26. https://doi.org/10.1016/j.swevo.2016.03.004

Abulaish, M., Wasi, N. A., & Sharma, S. (2024). The role of lifelong machine learning in bridging the gap between human and machine learning: A scientometric analysis. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *14*(2), e1526. https://doi.org/10.1002/widm.1526

Adelia, R., Suyanto, S., & Wisesty, U. N. (2019). Indonesian abstractive text summarization using bidirectional gated recurrent unit. *Procedia Computer Science*, *157*, 581–588. https://doi.org/10.1016/j.procs.2019.09.017

Ahmad, J., Zia, M. U., Naqvi, I. H., Chattha, J. N., Butt, F. A., Huang, T., & Xiang, W. (2024). Machine learning and blockchain technologies for cybersecurity in connected vehicles. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *14*(1), e1515. https://doi.org/10.1002/widm.1515

Ahmad, K., Pogorelov, K., Riegler, M., Conci, N., & Halvorsen, P. (2019). Social media and satellites: Disaster event detection, linking and summarization. *Multimedia Tools and Applications*, *78*, 2837–2875. https://doi.org/10.1007/s11042-018-5982-9

Akshaya, A. T., Shankaran, S., Thrupthi, H. M., & Mamatha, H. R. (2022). Natural language processing based cross lingual summarization. In Proceedings of the international conference on trends in electronics and informatics (pp. 1825–1829). https://doi.org/10.1109/ICOEI53556.2022.9776655

Alami Merrouni, Z., Frikh, B., & Ouhbi, B. (2023). EXABSUM: A new text summarization approach for generating extractive and abstractive summaries. *Journal of Big Data*, *10*(1), 163. https://doi.org/10.1186/s40537-023-00836-y

Aliakbarpour, H., Manzuri, M. T., & Rahmani, A. M. (2022). Improving the readability and saliency of abstractive text summarization using combination of deep neural networks equipped with auxiliary attention mechanism. *The Journal of Supercomputing*, *78*(2), 2528–2555. https://doi.org/10.1007/s11227-021-03950-x

Aliakbarpour, H., Manzuri, M. T., & Rahmani, A. M. (2024). Automatic text summarization using deep reinforced model coupling contextualized word representation and attention mechanism. *Multimedia Tools and Applications*, *83*(1), 733–762. https://doi.org/10.1007/s11042-023-15589-2

Alias, S., Mohammad, S. K., Hoon, G. K., & Ping, T. T. (2016). A Malay text summarizer using pattern-growth method with sentence compression rules. In 2016 third international conference on information retrieval and knowledge management (CAMP) (pp. 7–12). https://doi.org/10.1109/INFRKM.2016.7806326

Alias, S., Sainin, M. S., & Mohammad, S. K. (2021). A syntactic-based sentence validation technique for Malay text summarizer. *Journal of Information and Communication Technology*, *20*(3), 329–352. https://doi.org/10.32890/jict2021.20.3.3

Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). Text summarization techniques: A brief survey. *International Journal of Advanced Computer Science and Applications*, *8*(10), 397–405. https://doi.org/10.48550/arXiv.1707.02268

Almazroui, M., & Şen, Z. (2020). Trend analyses methodologies in hydro-meteorological records. *Earth Systems and Environment*, *4*, 713–738. https://doi.org/10.1007/s41748-020-00190-6

Alomari, A., Idris, N., Sabri, A. Q. M., & Alsmadi, I. (2022). Deep reinforcement and transfer learning for abstractive text summarization: A review. *Computer Speech & Language*, *71*, 101276. https://doi.org/10.1016/j.csl.2021.101276

Armenia, S., Franco, E., Iandolo, F., Maielli, G., & Vito, P. (2024). Zooming in and out the landscape: Artificial intelligence and system dynamics in business and management. *Technological Forecasting and Social Change*, *200*, 123131. https://doi.org/10.1016/j.techfore.2023.123131

Atri, Y. K., Goyal, V., & Chakraborty, T. (2023). Multi-document summarization using selective attention span and reinforcement learning. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *31*, 3457–3467. https://doi.org/10.1109/TASLP.2023.3316459

Babu, G. A., & Badugu, S. (2023). Deep learning based sequence to sequence model for abstractive telugu text summarization. *Multimedia Tools and Applications*, *82*(11), 17075–17096. https://doi.org/10.1007/s11042-022-14099-x

Banerjee, S., Mukherjee, S., Bandyopadhyay, S., & Pakray, P. (2023). An extract-then-abstract based method to generate disaster-news headlines using a DNN extractor followed by a transformer abstractor. *Information Processing & Management*, *60*(3), 103291. https://doi.org/10.1016/j.ipm.2023.103291

Bani-Almarjeh, M., & Kurdy, M. B. (2023). Arabic abstractive text summarization using RNN-based and transformer-based architectures. *Information Processing & Management*, *60*(2), 103227. https://doi.org/10.1016/j.ipm.2022.103227

Bannur, S., Hyland, S., Liu, Q., Perez-Garcia, F., Ilse, M., Castro, D. C., Boecking, B., Sharma, H., Bouzid, K., Thieme, A., Schwaighofer, A., Wetscherek, M., Lungren, M. P., Alvarez-Valle, A. N. J., & Oktay, O. (2023). Learning to exploit temporal structure for biomedical vision-language processing. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 15016–15027). https://doi.org/10.48550/arXiv.2301.04558

Bao, H., Dong, L., Wang, W., Yang, N., Piao, S., & Wei, F. (2023). Fine-tuning pretrained transformer encoders for sequence-to-sequence learning. *International Journal of Machine Learning and Cybernetics*, *1-18*, 1711–1728. https://doi.org/10.1007/s13042-023-01992-6

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. In Proceedings of the international AAAI conference on web and social media (Vol. 3, No. 1, pp. 361–362). https://doi.org/10.1609/icwsm.v3i1.13937

Baykara, B., & Güngör, T. (2023). Turkish abstractive text summarization using pretrained sequence-to-sequence models. *Natural Language Engineering*, *29*(5), 1275–1304. https://doi.org/10.1017/S1351324922000195

Beltagy, I., Peters, M. E., & Cohan, A. (2020). Longformer: The long-document transformer. arXiv Preprint 2004.05150. https://doi.org/10.48550/arXiv.2004.05150

Bhattacharjee, A., Hasan, T., Ahmad, W. U., Li, Y. F., Kang, Y. B., & Shahriyar, R. (2023). CrossSum: Beyond English-centric cross-lingual summarization for 1,500+ language pairs. In Annual meeting of the association of computational linguistics 2023 (pp. 2541–2564). https://doi.org/10.18653/v1/2023.acl-long.143

Bischof, J., & Airoldi, E. M. (2012). Summarizing topical content with word frequency and exclusivity. In Proceedings of the 29th international conference on machine learning (ICML-12) (pp. 201–208). https://doi.org/10.48550/arXiv.1206.4631

Bogers, M., Mention, A. L., & Barlatier, P. J. (2019). Social media and innovation: A systematic literature review and future research directions. *Technological Forecasting and Social Change*, *144*, 251–269. https://doi.org/10.1016/j.techfore.2018.10.007

Bui, N. D., Yu, Y., & Jiang, L. (2021). Self-supervised contrastive learning for code retrieval and summarization via semantic-preserving transformations. In Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval (pp. 511–521). https://doi.org/10.1145/3404835.3462840

Cai, X., & Li, W. (2013). Ranking through clustering: An integrated approach to multi-document summarization. *IEEE Transactions on Audio, Speech, and Language Processing*, *21*(7), 1424–1433. https://doi.org/10.1109/TASL.2013.2253098

Cai, X., Li, W., Ouyang, Y., & Yan, H. (2010). Simultaneous ranking and clustering of sentences: A reinforcement approach to multi-document summarization. In Proceedings of the 23rd international conference on computational linguistics (pp. 134–142). https://doi.org/10.5555/1873781.1873797

Cao, Y., Wan, X., Yao, J., & Yu, D. (2020). MultiSumm: Towards a unified model for multi-lingual abstractive summarization. In Proceedings of the AAAI conference on artificial intelligence (Vol. 34, No. 1, pp. 11–18). https://doi.org/10.1609/aaai.v34i01.5328

Carichon, F., Fettu, F., & Caporossi, G. (2023). Unsupervised update summarization of news events. *Pattern Recognition*, *144*, 109839. https://doi.org/10.1016/j.patcog.2023.109839

Chan, W., Zhou, X., Wang, W., & Chua, T. S. (2012). Community answer summarization for multi-sentence question with group l1 regularization. In Proceedings of the 50th annual meeting of the association for computational linguistics (Vol. 1: Long Papers) (pp. 582–591). https://doi.org/10.5555/2390524.2390606

Chen, X., Tao, X., Wang, F. L., & Xie, H. (2022). Global research on artificial intelligence-enhanced human electroencephalogram analysis. *Neural Computing and Applications*, *34*(14), 11295–11333. https://doi.org/10.1007/s00521-020-05588-x

Chen, X., Xie, H., & Hwang, G. J. (2020). A multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, institutions, and researchers. *Computers and Education: Artificial Intelligence*, *1*, 100005. https://doi.org/10.1016/j.caeai.2020.100005

Chen, X., Xie, H., Li, Z., Cheng, G., Leng, M., & Wang, F. L. (2023). Information fusion and artificial intelligence for smart healthcare: A bibliometric study. *Information Processing & Management*, *60*(1), 103113. https://doi.org/10.1016/j.ipm.2022.103113

Chen, X., Zou, D., Cheng, G., & Xie, H. (2020). Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A retrospective of all volumes of computer & education. *Computers & Education*, *151*, 103855. https://doi.org/10.1016/j.compedu.2020.103855

Chen, Y., Liu, Y., Xu, R., Yang, Z., Zhu, C., Zeng, M., & Zhang, Y. (2023). UniSumm and SummZoo: Unified model and diverse benchmark for few-shot summarization. In Proceedings of the 61st annual meeting of the association for computational linguistics (Vol. 1: Long Papers) (pp. 12833–12855). https://doi.org/10.18653/v1/2023.acl-long.718

Chen, Y. C., & Bansal, M. (2018). Fast abstractive summarization with reinforce-selected sentence rewriting. In Proceedings of the 56th annual meeting of the association for computational linguistics (Vol. 1: Long Papers) (pp. 675–686). https://doi.org/10.18653/v1/P18-1063

Chen, Y. S., & Shuai, H. H. (2021). Meta-transfer learning for low-resource abstractive summarization. In Proceedings of the AAAI conference on artificial intelligence (Vol. 35, No. 14, pp. 12692–12700). https://doi.org/10.1609/aaai.v35i14.17503

Cheng, J., & Lapata, M. (2016). Neural summarization by extracting sentences and words. In Proceedings of the 54th annual meeting of the association for computational linguistics (Vol. 1: Long Papers) (pp. 484–494). https://doi.org/10.18653/v1/P16-1046

Cheng, Z., & Guo, S. (2022). Automatic text summarization for public health WeChat official accounts platform base on improved TextRank. *Journal of Environmental and Public Health*, *2022*, 1166989. https://doi.org/10.1155/2022/1166989

Christian, H., Agus, M. P., & Suhartono, D. (2016). Single document automatic text summarization using term frequency-inverse document frequency (TF-IDF). *ComTech: Computer, Mathematics and Engineering Applications*, *7*(4), 285–294. https://doi.org/10.21512/comtech.v7i4.3746

Dalmia, S., Okhonko, D., Lewis, M., Edunov, S., Watanabe, S., Metze, F., Zettlemoyer, L., & Mohamed, A. (2023). Legonn: building modular encoder-decoder models. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *31*, 3112–3126. https://doi.org/10.1109/TASLP.2023.3296019

Debnath, D., Das, R., & Pakray, P. (2021). Extractive single document summarization using multi-objective modified cat swarm optimization approach: ESDS-MCSO. *Neural Computing and Applications*, 1–16. https://doi.org/10.1007/s00521-021-06337-4

Demir, E. K. (2021). The role of social capital for teacher professional learning and student achievement: A systematic literature review. *Educational Research Review*, *33*, 100391. https://doi.org/10.1016/j.edurev.2021.100391

Deng, Y., Zhang, W., & Lam, W. (2020). Multi-hop inference for question-driven summarization. In Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP) (pp. 6734–6744). https://doi.org/10.18653/v1/2020.emnlp-main.547

Deng, Y., Zhang, W., Xu, W., Shen, Y., & Lam, W. (2023). Nonfactoid question answering as query-focused summarization with graph-enhanced multihop inference. *IEEE Transactions on Neural Networks and Learning Systems*. https://doi.org/10.1109/TNNLS.2023.3258413

Devarakonda, M., Zhang, D., Tsou, C. H., & Bornea, M. (2014). Problem-oriented patient record summary: An early report on a Watson application. In 2014 IEEE 16th international conference on e-health networking, applications and services (Healthcom) (pp. 281–286). https://doi.org/10.1109/HealthCom.2014.7001855

Dong, L., Yang, N., Wang, W., Wei, F., Liu, X., Wang, Y., Gao, J., Zhou, M., & Hon, H. W. (2019). Unified language model pre-training for natural language understanding and generation. In Proceedings of the 33rd international conference on neural information processing systems (pp. 13063–13075). https://doi.org/10.5555/3454287.3455457

Dong, M., Li, F., & Chang, H. (2023). Trends and hotspots in critical thinking research over the past two decades: Insights from a bibliometric analysis. *Heliyon*, *32*, e16934. https://doi.org/10.1016/j.heliyon.2023.e16934

Donthu, N., Kumar, S., Pandey, N., Pandey, N., & Mishra, A. (2021). Mapping the electronic word-of-mouth (eWOM) research: A systematic review and bibliometric analysis. *Journal of Business Research*, *135*, 758–773. https://doi.org/10.1016/j.jbusres.2021.07.015

Dwivedi, Y. K., Sharma, A., Rana, N. P., Giannakis, M., Goel, P., & Dutot, V. (2023). Evolution of artificial intelligence research in technological forecasting and social change: Research topics, trends, and future directions. *Technological Forecasting and Social Change*, *192*, 122579. https://doi.org/10.1016/j.techfore.2023.122579

El-Kassas, W. S., Salama, C. R., Rafea, A. A., & Mohamed, H. K. (2021). Automatic text summarization: A comprehensive survey. *Expert Systems with Applications*, *165*, 113679. https://doi.org/10.1016/j.eswa.2020.113679

Elsahar, H., Coavoux, M., Rozen, J., & Gallé, M. (2021). Self-supervised and controlled multi-document opinion summarization. In Proceedings of the 16th conference of the european chapter of the association for computational linguistics: Main volume (pp. 1646–1662). https://doi.org/10.18653/v1/2021.eacl-main.141

Fabbri, A. R., Han, S., Li, H., Li, H., Ghazvininejad, M., Joty, S., Radev, D., & Mehdad, Y. (2021). Improving zero and few-shot abstractive summarization with intermediate fine-tuning and data augmentation. In Proceedings of the 2021 conference of the north american chapter of the association for computational linguistics: Human language technologies (pp. 704–717). https://doi.org/10.18653/v1/2021.naacl-main.57

Fauzi, M. A. (2022). E-learning in higher education institutions during COVID-19 pandemic: Current and future trends through bibliometric analysis. *Heliyon*, *8*(5), e09433. https://doi.org/10.1016/j.heliyon.2022.e09433

Feblowitz, J. C., Wright, A., Singh, H., Samal, L., & Sittig, D. F. (2011). Summarization of clinical information: A conceptual model. *Journal of Biomedical Informatics*, *44*(4), 688–699. https://doi.org/10.1016/j.jbi.2011.03.008

Feinerer, I., Hornik, K., & Feinerer, M. I. (2015). Package 'tm'. https://cran.r-project.org/web/packages/tm/tm.pdf

Frisoni, G., Italiani, P., Salvatori, S., & Moro, G. (2023). Cogito ergo summ: Abstractive summarization of biomedical papers via semantic parsing graphs and consistency rewards. In Proceedings of the AAAI conference on artificial intelligence (Vol. 37, No. 11, pp. 12781–12789). https://doi.org/10.1609/aaai.v37i11.26503

Gambhir, M., & Gupta, V. (2017). Recent automatic text summarization techniques: a survey. *Artificial Intelligence Review*, *47*, 1–66. https://doi.org/10.1007/s10462-016-9475-9

Gambhir, M., & Gupta, V. (2022). Deep learning-based extractive text summarization with word-level attention mechanism. *Multimedia Tools and Applications*, *81*(15), 20829–20852. https://doi.org/10.1007/s11042-022-12729-y

Gao, Y., Meyer, C. M., & Gurevych, I. (2018). APRIL: Interactively learning to summarise by combining active preference learning and reinforcement learning. In Proceedings of the 2018 conference on empirical methods in natural language processing (pp. 4120–4130). https://doi.org/10.18653/v1/D18-1445

Gao, Y., Zhao, W., & Eger, S. (2020). SUPERT: Towards new frontiers in unsupervised evaluation metrics for multi-document summarization. In Proceedings of the 58th annual meeting of the association for computational linguistics (pp. 1347–1354). https://doi.org/10.18653/v1/2020.acl-main.124

Gehrmann, S., Deng, Y., & Rush, A. M. (2018). Bottom-up abstractive summarization. In Proceedings of the 2018 conference on empirical methods in natural language processing (pp. 4098–4109). https://doi.org/10.18653/v1/D18-1443

Ghadimi, A., & Beigy, H. (2022). Hybrid multi-document summarization using pre-trained language models. *Expert Systems with Applications*, *192*, 116292. https://doi.org/10.1016/j.eswa.2021.116292

Ghadimi, A., & Beigy, H. (2023). SGCSumm: An extractive multi-document summarization method based on pre-trained language model, submodularity, and graph convolutional neural networks. *Expert Systems with Applications*, *215*, 119308. https://doi.org/10.1016/j.eswa.2022.119308

Ghalehtaki, R. A., Khotanlou, H., & Esmaeilpour, M. (2014). A combinational method of fuzzy, particle swarm optimization and cellular learning automata for text summarization. In 2014 Iranian conference on intelligent systems (ICIS) (pp. 1–6). https://doi.org/10.1109/IranianCIS.2014.6802577

Giarelis, N., Mastrokostas, C., & Karacapilidis, N. (2023). Abstractive vs. extractive summarization: An experimental review. *Applied Sciences*, *13*(13), 7620. https://doi.org/10.3390/app13137620

Givchi, A., Ramezani, R., & Baraani-Dastjerdi, A. (2022). Graph-based abstractive biomedical text summarization. *Journal of Biomedical Informatics*, *132*, 104099. https://doi.org/10.1016/j.jbi.2022.104099

Goodwin, T., Savery, M., & Demner-Fushman, D. (2020). Flight of the PEGASUS? Comparing transformers on few-shot and zero-shot multi-document abstractive summarization. In Proceedings of the 28th international conference on computational linguistics (pp. 5640–5646). https://doi.org/10.18653/v1/2020.coling-main.494

Goyal, R., Kumar, P., & Singh, V. P. (2023). A systematic survey on automated text generation tools and techniques: Application, evaluation, and challenges. *Multimedia Tools and Applications*, *82*, 43089–43144. https://doi.org/10.1007/s11042-023-15224-0

Grail, Q., Perez, J., & Gaussier, E. (2021). Globalizing BERT-based transformer architectures for long document summarization. In Proceedings of the 16th conference of the european chapter of the association for computational linguistics: Main volume (pp. 1792–1810). https://doi.org/10.18653/v1/2021.eacl-main.154

Gulden, C., Kirchner, M., Schüttler, C., Hinderer, M., Kampf, M., Prokosch, H.-U., & Toddenroth, D. (2019). Extractive summarization of clinical trial descriptions. *International Journal of Medical Informatics*, *129*, 114–121. https://doi.org/10.1016/j.ijmedinf.2019.05.019

Guo, Q., Huang, J., Xiong, N., & Wang, P. (2019). MS-pointer network: Abstractive text summary based on multi-head self-attention. *IEEE Access*, *7*, 138603–138613. https://doi.org/10.1109/ACCESS.2019.2941964

Gupta, S., & Gupta, S. K. (2019). Abstractive summarization: An overview of the state of the art. *Expert Systems with Applications*, *121*, 49–65. https://doi.org/10.1016/j.eswa.2018.12.011

Gurcan, F., Cagiltay, N. E., & Cagiltay, K. (2021). Mapping human–computer interaction research themes and trends from its existence to today: A topic modeling-based review of past 60 years. *International Journal of Human–Computer Interaction*, *37*(3), 267–280. https://doi.org/10.1080/10447318.2020.1819668

Han, X., Lv, T., Hu, Z., Wang, X., & Wang, C. (2016). Text summarization using framenet-based semantic graph model. *Scientific Programming*, *2016*, 5130603. https://doi.org/10.1155/2016/5130603

Iqbal, W., Malik, W. I., Bukhari, F., Almustafa, K. M., & Nawaz, Z. (2021). Big data full-text search index minimization using text summarization. *Information Technology and Control*, *50*(2), 375–389. https://doi.org/10.5755/j01.itc.50.2.25470

Iqbal, W., Qadir, J., Tyson, G., Mian, A. N., Hassan, S. U., & Crowcroft, J. (2019). A bibliometric analysis of publications in computer networking research. *Scientometrics*, *119*, 1121–1155. https://doi.org/10.1007/s11192-019-03086-z

Jangra, A., Mukherjee, S., Jatowt, A., Saha, S., & Hasanuzzaman, M. (2023). A survey on multi-modal summarization. *ACM Computing Surveys*, *55*(13), 1–36. https://doi.org/10.1145/3584700

Jia, R., Zhang, X., Cao, Y., Lin, Z., Wang, S., & Wei, F. (2022). Neural label search for zero-shot multi-lingual extractive summarization. In Proceedings of the 60th annual meeting of the association for computational linguistics (Vol. 1: Long Papers) (pp. 561–570). https://doi.org/10.18653/v1/2022.acl-long.42

Jiang, H., Xu, S., & Lau, F. C. M. (2011). Capturing user reading behaviors for personalized document summarization. In Proceedings of the 16th international conference on intelligent user interfaces (pp. 355–358). https://doi.org/10.1145/1943403.1943464

Jiang, Y., & Wang, Y. (2023). Topic-aware hierarchical multi-attention network for text classification. *International Journal of Machine Learning and Cybernetics*, *14*(5), 1863–1875. https://doi.org/10.1007/s13042-022-01734-0

Jiang, Z., Cai, X., Yang, L., Gao, D., Zhao, W., Han, J., Jun, L., Dinggang, S., & Liu, T. (2023). Learning to summarize Chinese radiology findings with a pre-trained encoder. *IEEE Transactions on Biomedical Engineering*, *70*, 3277–3287. https://doi.org/10.1109/TBME.2023.3280987

Joshi, A., Fidalgo, E., Alegre, E., & Fernández-Robles, L. (2023). DeepSumm: Exploiting topic models and sequence to sequence networks for extractive text summarization. *Expert Systems with Applications*, *211*, 118442. https://doi.org/10.1016/j.eswa.2022.118442

Kanwal, N., & Rizzo, G. (2022). Attention-based clinical note summarization. In Proceedings of the 37th ACM/SIGAPP symposium on applied computing (pp. 813–820). https://doi.org/10.1145/3477314.3507256

Kim, B., Kang, S., & Lee, S. (2019). A weighted pagerank-based bug report summarization method using bug report relationships. *Applied Sciences*, *9*(24), 5427. https://doi.org/10.3390/app9245427

Koutras, P., Zlatintsi, A., Iosif, E., Katsamanis, A., Maragos, P., & Potamianos, A. (2015). Predicting audio-visual salient events based on visual, audio and text modalities for movie summarization. In 2015 IEEE international conference on image processing (ICIP) (pp. 4361–4365). https://doi.org/10.1109/ICIP.2015.7351630

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

Kumar, A., Seth, S., Gupta, S., & Maini, S. (2022). Sentic computing for aspect-based opinion summarization using multi-head attention with feature pooled pointer generator network. *Cognitive Computation*, *14*(1), 130–148. https://doi.org/10.1007/s12559-021-09835-8

Kumar, K. V., Yadav, D., & Sharma, A. (2015). Graph based technique for hindi text summarization. In Proceedings of second international conference of information systems design and intelligent applications (Vol. 1, pp. 301–310). Springer India. https://doi.org/10.1007/978-81-322-2250-7_29

Kumar, S., & Solanki, A. (2023). An abstractive text summarization technique using transformer model with self-attention mechanism. *Neural Computing and Applications*, *35*, 18603–18622. https://doi.org/10.1007/s00521-023-08687-7

Lau, J. H., Newman, D., & Baldwin, T. (2014). Machine reading tea leaves: Automatically evaluating topic coherence and topic model quality. In Proceedings of the 14th conference of the European chapter of the association for computational linguistics (pp. 530–539).

Lee, C.-H., Yang, H.-C., Chen, Y. J., & Chuang, Y.-L. (2021). Event monitoring and intelligence gathering using Twitter based real-time event summarization and pre-trained model techniques. *Applied Sciences*, *11*(22), 10596. https://doi.org/10.3390/app112210596

Li, D., Mei, H., Shen, Y., Su, S., Zhang, W., Wang, J., Zu, M., & Chen, W. (2018). ECharts: A declarative framework for rapid construction of web-based visualization. *Visual Informatics*, *2*(2), 136–146. https://doi.org/10.1016/j.visinf.2018.04.011

Li, H., Peng, Q., Mou, X., Wang, Y., Zeng, Z., & Bashir, M. F. (2023). Abstractive financial news summarization via transformer-BiLSTM encoder and graph attention-based decoder. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *31*, 3190–3205. https://doi.org/10.1109/TASLP.2023.3304473

Li, H., Zhu, J., Ma, C., Zhang, J., & Zong, C. (2017). Multi-modal summarization for asynchronous collection of text, image, audio and video. In Proceedings of the 2017 conference on empirical methods in natural language processing (pp. 1092–1102). https://doi.org/10.18653/v1/D17-1114

Li, J. (2023). Information avoidance in the age of COVID-19: A meta-analysis. *Information Processing & Management*, *60*(1), 103163. https://doi.org/10.1016/j.ipm.2022.103163

Li, J., Zhang, C., Chen, X., Cao, Y., Liao, P., & Zhang, P. (2019). Abstractive text summarization with multi-head attention. In 2019 international joint conference on neural networks (pp. 1–8). https://doi.org/10.1109/IJCNN.2019.8851885

Li, L., & Li, T. (2013). An empirical study of ontology-based multi-document summarization in disaster management. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *44*(2), 162–171. https://doi.org/10.1109/TSMCC.2013.2258335

Li, Q., Peng, H., Li, J., Wu, J., Ning, Y., Wang, L., Yu, P. S., & Wang, Z. (2021). Reinforcement learning-based dialogue guided event extraction to exploit argument relations. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *30*, 520–533. https://doi.org/10.1109/TASLP.2021.3138670

Li, S., & Xu, J. (2023). HierMDS: A hierarchical multi-document summarization model with global–local document dependencies. *Neural Computing and Applications*, *35*, 18553–18570. https://doi.org/10.1007/s00521-023-08680-0

Li, Y., Huang, Y., Huang, W., & Wang, W. (2024). A global and local information extraction model incorporating selection mechanism for abstractive text summarization. *Multimedia Tools and Applications*, *83*(2), 4859–4886. https://doi.org/10.1007/s11042-023-15274-4

Liang, X., & Liu, A. M. (2018). The evolution of government sponsored collaboration network and its impact on innovation: A bibliometric analysis in the Chinese solar PV sector. *Research Policy*, *47*(7), 1295–1308. https://doi.org/10.1016/j.respol.2018.04.012

Lin, H., & Bilmes, J. (2011). A class of submodular functions for document summarization. In Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies (pp. 510–520). https://doi.org/10.5555/2002472.2002537

Liu, C.-Y., Chen, M.-S., & Tseng, C.-Y. (2015). Incrests: Towards real-time incremental short text summarization on comment streams from social network services. *IEEE Transactions on Knowledge and Data Engineering*, *27*(11), 2986–3000. https://doi.org/10.1109/TKDE.2015.2405553

Liu, P. J., Saleh, M., Pot, E., Goodrich, B., Sepassi, R., Kaiser, L., & Shazeer, N. (2018). Generating wikipedia by summarizing long sequences. arXiv Preprint: 1801.10198. https://doi.org/10.48550/arXiv.1801.10198

Liu, S., Cao, J., Yang, R., & Wen, Z. (2022). Key phrase aware transformer for abstractive summarization. *Information Processing & Management*, *59*(3), 102913. https://doi.org/10.1016/j.ipm.2022.102913

Liu, S., Yang, L., & Cai, X. (2022). SEASum: Syntax-enriched abstractive summarization. *Expert Systems with Applications*, *199*, 116819. https://doi.org/10.1016/j.eswa.2022.116819

Liu, Y., & Lapata, M. (2019a). Hierarchical transformers for multi-document summarization. In Proceedings of the 57th annual meeting of the association for computational linguistics (pp. 5070–5081). https://doi.org/10.18653/v1/P19-1500

Liu, Y., & Lapata, M. (2019b). Text summarization with pretrained encoders. In Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP) (pp. 3730–3740). https://doi.org/10.18653/v1/D19-1387

Loper, E., & Bird, S. (2002). NLTK: The natural language toolkit. In Proceedings of the ACL-02 workshop on effective tools and methodologies for teaching natural language processing and computational linguistics-volume 1 (pp. 63–70). https://doi.org/10.48550/arXiv.cs/0205028

Lu, Y., Dong, Y., & Charlin, L. (2020). Multi-XScience: a large-scale dataset for extreme multi-document summarization of scientific articles. In Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP) (pp. 8068–8074). https://doi.org/10.18653/v1/2020.emnlp-main.648

Luhn, H. P. (1958). The automatic creation of literature abstracts. *IBM Journal of Research and Development*, *2*(2), 159–165. https://doi.org/10.1147/rd.22.0159

Ma, X., Keung, J. W., Yu, X., Zou, H., Zhang, J., & Li, Y. (2023). AttSum: A deep attention-based summarization model for bug report title generation. *IEEE Transactions on Reliability*, *72*, 1663–1677. https://doi.org/10.1109/TR.2023.3236404

Mahajani, A., Pandya, V., Maria, I., & Sharma, D. (2019). A comprehensive survey on extractive and abstractive techniques for text summarization. *Ambient Communications and Computer Systems*, *904*, 339–351. https://doi.org/10.1007/978-981-13-5934-7_31

Mani, S., Catherine, R., Sinha, V. S., & Dubey, A. (2012). Ausum: Approach for unsupervised bug report summarization. In Proceedings of the ACM SIGSOFT 20th international symposium on the foundations of software engineering (pp. 1–11). https://doi.org/10.1145/2393596.2393607

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica: Journal of the Econometric Society*, *13*, 245–259. https://doi.org/10.2307/1907187

Marujo, L., Ling, W., Ribeiro, R., Gershman, A., Carbonell, J., de Matos, D. M., & Neto, J. P. (2016). Exploring events and distributed representations of text in multi-document summarization. *Knowledge-Based Systems*, *94*, 33–42. https://doi.org/10.1016/j.knosys.2015.11.005

McCreadie, R., Macdonald, C., & Ounis, I. (2018). Automatic ground truth expansion for timeline evaluation. In the 41st international ACM SIGIR conference on research & development in information retrieval (pp. 685–694). https://doi.org/10.1145/3209978.3210034

Meena, P., Kumar, H., & Yadav, S. K. (2023). A review on video summarization techniques. *Engineering Applications of Artificial Intelligence*, *118*, 105667. https://doi.org/10.1016/j.engappai.2022.105667

Mimno, D., Wallach, H., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. In Proceedings of the 2011 conference on empirical methods in natural language processing (pp. 262–272). https://aclanthology.org/D11-1024

Mishra, R., Bian, J., Fiszman, M., Weir, C. R., Jonnalagadda, S., Mostafa, J., & Del Fiol, G. (2014). Text summarization in the biomedical domain: A systematic review of recent research. *Journal of Biomedical Informatics*, *52*, 457–467. https://doi.org/10.1016/j.jbi.2014.06.009

Mohamed, M., & Oussalah, M. (2019). SRL-ESA-TextSum: A text summarization approach based on semantic role labeling and explicit semantic analysis. *Information Processing & Management*, *56*(4), 1356–1372. https://doi.org/10.1016/j.ipm.2019.04.003

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Moratanch, N., & Chitrakala, S. (2017). A survey on extractive text summarization. In 2017 international conference on computer, communication and signal processing (ICCCSP) (pp. 1–6). https://doi.org/10.1109/ICCCSP.2017.7944061

Mrinalini, K., Nagarajan, T., & Vijayalakshmi, P. (2018). Pause-based phrase extraction and effective OOV handling for low-resource machine translation systems. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, *18*(2), 1–22. https://doi.org/10.1145/3265751

Mu, L., Jin, P., Zhang, Y., Zhong, H., & Zhao, J. (2023). Synonym recognition from short texts: A self-supervised learning approach. *Expert Systems with Applications*, *224*, 119966. https://doi.org/10.1016/j.eswa.2023.119966

Musto, C., Rossiello, G., de Gemmis, M., Lops, P., & Semeraro, G. (2019). Combining text summarization and aspect-based sentiment analysis of users' reviews to justify recommendations. In Proceedings of the 13th ACM conference on recommender systems (pp. 383–387). https://doi.org/10.1145/3298689.3347024

Mutlu, B., & Sezer, E. A. (2023). Enhanced sentence representation for extractive text summarization: Investigating the syntactic and semantic features and their contribution to sentence scoring. *Expert Systems with Applications*, *227*, 120302. https://doi.org/10.1016/j.eswa.2023.120302

Nallapati, R., Zhai, F., & Zhou, B. (2017). SummaRuNNer: A recurrent neural network based sequence model for extractive summarization of documents. In Proceedings of the thirty-first AAAI conference on artificial intelligence (pp. 3075–3081). https://doi.org/10.1609/aaai.v31i1.10958

Nallapati, R., Zhou, B., dos Santos, C., Gůlçehre, Ç., & Xiang, B. (2016). Abstractive text summarization using sequence-to-sequence RNNs and beyond. In Proceedings of the 20th SIGNLL conference on computational natural language learning (pp. 280–290). https://doi.org/10.18653/v1/K16-1028

Narayan, S., Cohen, S. B., & Lapata, M. (2018a). Don't give me the details, just the summary! Topic-aware convolutional neural networks for extreme summarization. In Proceedings of the 2018 conference on empirical methods in natural language processing (pp. 1797–1807). https://doi.org/10.18653/v1/D18-1206

Narayan, S., Cohen, S. B., & Lapata, M. (2018b). Ranking sentences for extractive summarization with reinforcement learning. In 16th annual conference of the North American chapter of the association for computational linguistics: Human language technologies (pp. 1747–1759). https://doi.org/10.18653/v1/N18-1158

Ouyang, J., Song, B., & McKeown, K. (2019). A robust abstractive system for cross-lingual summarization. In Proceedings of the 2019 conference of the north american chapter of the association for computational linguistics: Human language technologies (Vol. 1, pp. 2025–2031). https://doi.org/10.18653/v1/N19-1204

Parnell, J., Unanue, I. J., & Piccardi, M. (2021). RewardsOfSum: Exploring reinforcement learning rewards for summarisation. In Proceedings of the 5th workshop on structured prediction for NLP (SPNLP 2021) (pp. 1–11). https://doi.org/10.18653/v1/2021.spnlp-1.1

Patel, D., Shah, S., & Chhinkaniwala, H. (2019). Fuzzy logic based multi document summarization with improved sentence scoring and redundancy removal technique. *Expert Systems with Applications*, *134*, 167–177. https://doi.org/10.1016/j.eswa.2019.05.045

Paulus, R., Xiong, C., & Socher, R. (2017). A deep reinforced model for abstractive summarization. arXiv Preprint:1705.04304. https://doi.org/10.48550/arXiv.1705.04304

Pian, W., Chi, J., & Ma, F. (2021). The causes, impacts and countermeasures of COVID-19 "Infodemic": A systematic review using narrative synthesis. *Information Processing & Management*, *58*(6), 102713. https://doi.org/10.1016/j.ipm.2021.102713

Plaza, L., Díaz, A., & Gervás, P. (2011). A semantic graph-based approach to biomedical summarisation. *Artificial Intelligence in Medicine*, *53*(1), 1–14. https://doi.org/10.1016/j.artmed.2011.06.005

Pohlert, T. (2016). Package "trend". https://cran.r-project.org/web//packages/trend/vignettes/trend.pdf

Pontes, E. L., Huet, S., Torres-Moreno, J.-M., & Linhares, A. C. (2020). Compressive approaches for cross-language multi-document summarization. *Data & Knowledge Engineering*, *125*, 101763. https://doi.org/10.1016/j.datak.2019.101763

Porntrakoon, P., Moemeng, C., & Santiprabhob, P. (2021). Text summarization for Thai food reviews using simplified sentiment analysis. In 2021 18th international joint conference on computer science and software engineering (JCSSE) (pp. 1–5). https://doi.org/10.1109/JCSSE53117.2021.9493839

Pradhan, T., Bhatia, C., Kumar, P., & Pal, S. (2021). A deep neural architecture based meta-review generation and final decision prediction of a scholarly article. *Neurocomputing*, *428*, 218–238. https://doi.org/10.1016/j.neucom.2020.11.004

Punj, N., Ahmi, A., Tanwar, A., & Rahim, S. A. (2023). Mapping the field of green manufacturing: A bibliometric review of the literature and research frontiers. *Journal of Cleaner Production*, *423*, 138729. https://doi.org/10.1016/j.jclepro.2023.138729

Qin, Z., Yang, S., & Zhong, Y. (2024). Hierarchically gated recurrent neural network for sequence modeling. *Advances in Neural Information Processing Systems*, *36*, 1–20. https://doi.org/10.48550/arXiv.2311.04823

Rahman, A. U., Alsenani, Y., Zafar, A., Ullah, K., Rabie, K., & Shongwe, T. (2024). Enhancing heart disease prediction using a self-attention-based transformer model. *Scientific Reports*, *14*(1), 514. https://doi.org/10.1038/s41598-024-51184-7

Rajan, R. P., & Jose, D. V. (2023). Text summarization using residual-based temporal attention convolutional neural network. *International Journal of Information Technology*, *1-19*. https://doi.org/10.1007/s41870-023-01581-4

Ramachandran, L., & Foltz, P. (2015). Generating reference texts for short answer scoring using graph-based summarization. In Proceedings of the tenth workshop on innovative use of nlp for building educational applications (pp. 207–212). https://doi.org/10.3115/v1/W15-0624

Ramesh Kashyap, A., Yang, Y., & Kan, M. Y. (2023). Scientific document processing: Challenges for modern learning methods. *International Journal on Digital Libraries*, *24*, 283–309. https://doi.org/10.1007/s00799-023-00352-7

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*, *91*, 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., Albertson, B., & 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

Sabha, A., & Selwal, A. (2023). Data-driven enabled approaches for criteria-based video summarization: A comprehensive survey, taxonomy, and future directions. *Multimedia Tools and Applications*, *82*, 32635–32709. https://doi.org/10.1007/s11042-023-14925-w

Saggion, H., & Poibeau, T. (2013). Automatic text summarization: Past, present and future. In T. Poibeau, H. Saggion, J. Piskorski, & R. Yangarber (Eds.), *Multi-source, multilingual information extraction and summarization. Theory and applications of natural language processing* (pp. 3–21). Springer. https://doi.org/10.1007/978-3-642-28569-1_1

Sanabria, M., Precioso, F., & Menguy, T. (2021). Hierarchical multimodal attention for deep video summarization. In 2020 25th international conference on pattern recognition (ICPR) (pp. 7977–7984). https://doi.org/10.1109/ICPR48806.2021.9413097

Searle, T., Ibrahim, Z., Teo, J., & Dobson, R. J. (2023). Discharge summary hospital course summarisation of in patient electronic health record text with clinical concept guided deep pre-trained transformer models. *Journal of Biomedical Informatics*, *141*, 104358. https://doi.org/10.1016/j.jbi.2023.104358

Shams, R., & Mercer, R. E. (2015). Summary sentence classification using stylometry. In 2015 IEEE 14th international conference on machine learning and applications (ICMLA) (pp. 1220–1227). https://doi.org/10.1109/ICMLA.2015.181

Sharma, A., Rana, N. P., & Nunkoo, R. (2021). Fifty years of information management research: A conceptual structure analysis using structural topic modeling. *International Journal of Information Management*, *58*, 102316. https://doi.org/10.1016/j.ijinfomgt.2021.102316

Shi, T., Keneshloo, Y., Ramakrishnan, N., & Reddy, C. K. (2021). Neural abstractive text summarization with sequence-to-sequence models. *ACM Transactions on Data Science*, *2*(1), 1–37. https://doi.org/10.48550/arXiv.1812.02303

Shi, X. (2023). MCLS: A large-scale multimodal cross-lingual summarization dataset. In China national conference on chinese computational linguistics (pp. 273–288). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-6207-5_17

Singh, G., Mittal, N., & Chouhan, S. S. (2024). A deep learning framework for multi-document summarization using LSTM with improved dingo optimizer (IDO). *Multimedia Tools and Applications*, *1-23*. https://doi.org/10.1007/s11042-024-18248-2

Song, Y., Liang, N., Guo, Q., Dai, J., Bai, J., & He, F. (2023). MeshCLIP: Efficient cross-modal information processing for 3D mesh data in zero/few-shot learning. *Information Processing & Management*, *60*(6), 103497. https://doi.org/10.1016/j.ipm.2023.103497

Soni, S., Chouhan, S. S., & Rathore, S. S. (2023). TextConvoNet: A convolutional neural network based architecture for text classification. *Applied Intelligence*, *53*(11), 14249–14268. https://doi.org/10.1007/s10489-022-04221-9

Sun, F., & Tian, X. (2022). Lecture video automatic summarization system based on DBNet and Kalman filtering. *Mathematical Problems in Engineering*, *2022*, 5303503. https://doi.org/10.1155/2022/5303503

Swetha, G., & Kumar, S. P. (2023). A hierarchical framework based on transformer technology to achieve factual consistent and non-redundant abstractive text summarization. *Multimedia Tools and Applications*, *1-22*. https://doi.org/10.1007/s11042-023-17426-y

Takeshita, S., Green, T., Friedrich, N., Eckert, K., & Ponzetto, S. P. (2023). Cross-lingual extreme summarization of scholarly documents. *International Journal on Digital Libraries*, *1-23*. https://doi.org/10.1007/s00799-023-00373-2

Uc-Cetina, V., Navarro-Guerrero, N., Martin-Gonzalez, A., Weber, C., & Wermter, S. (2023). Survey on reinforcement learning for language processing. *Artificial Intelligence Review*, *56*(2), 1543–1575. https://doi.org/10.1007/s10462-022-10205-5

Van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3

Wahab, M. H. H., Hamid, N. A. W. A., Subramaniam, S., Latip, R., & Othman, M. (2024). Decomposition–based multi-objective differential evolution for extractive multi-document automatic text summarization. *Applied Soft Computing*, *151*, 110994. https://doi.org/10.1016/j.asoc.2023.110994

Wamba, S. F., Queiroz, M. M., & Hamzi, L. (2023). A bibliometric and multi-disciplinary quasi-systematic analysis of social robots: Past, future, and insights of human-robot interaction. *Technological Forecasting and Social Change*, *197*, 122912. https://doi.org/10.1016/j.techfore.2023.122912

Wan, D., & Bansal, M. (2022). FactPEGASUS: Factuality-aware pre-training and fine-tuning for abstractive summarization. In Proceedings of the 2022 conference of the north american chapter of the association for computational linguistics: Human language technologies (pp. 1010–1028). https://doi.org/10.18653/v1/2022.naacl-main.74

Wang, H., Wang, X., Xiong, W., Yu, M., Guo, X., Chang, S., & Wang, W. Y. (2019). Self-supervised learning for contextualized extractive summarization. In Proceedings of the 57th annual meeting of the association for computational linguistics (pp. 2221–2227). https://doi.org/10.18653/v1/P19-1214

Wang, H., & Zhou, G. (2010). Topic-driven multi-document summarization. In 2010 international conference on Asian language processing (pp. 195–198). https://doi.org/10.1109/IALP.2010.26

Wang, P., Li, S., Liu, S., Tang, J., & Wang, T. (2023). Plan and generate: Explicit and implicit variational augmentation for multi-document summarization of scientific articles. *Information Processing & Management*, *60*(4), 103409. https://doi.org/10.1016/j.ipm.2023.103409

Wang, S., Zhao, X., Li, B., Ge, B., & Tang, D. (2017). Integrating extractive and abstractive models for long text summarization. In 2017 IEEE international congress on big data (BigData congress) (pp. 305–312). https://doi.org/10.1109/BigDataCongress.2017.46

Wang, X., Yoshida, Y., Hirao, T., Sudoh, K., & Nagata, M. (2015). Summarization based on task-oriented discourse parsing. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *23*(8), 1358–1367. https://doi.org/10.1109/TASLP.2015.2432573

Wei, R., Huang, H., & Gao, Y. (2019). Sharing pre-trained BERT decoder for a hybrid summarization. In Proceedings of 18th chinese computational linguistics (Vol. 18, pp. 169–180). https://doi.org/10.1007/978-3-030-32381-3_14

West, P., Holtzman, A., Buys, J., & Choi, Y. (2019). BottleSum: Unsupervised and self-supervised sentence summarization using the information bottleneck principle. In Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP) (pp. 3752–3761). https://doi.org/10.18653/v1/D19-1389

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., & van den Brand, T. (2016). ggplot2: create elegant data visualisations using the grammar of graphics. R package version 3.2. 1. Stata Software Package. https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf

Widyassari, A. P., Rustad, S., Shidik, G. F., Noersasongko, E., Syukur, A., & Affandy, A. (2022). Review of automatic text summarization techniques & methods. *Journal of King Saud University-Computer and Information Sciences*, *34*(4), 1029–1046. https://doi.org/10.1016/j.jksuci.2020.05.006

Wu, H. (2024). Dilated convolution for enhanced extractive summarization: A GAN-based approach with BERT word embedding. *Journal of Intelligent & Fuzzy Systems*, *46*(2), 4777–4790. https://doi.org/10.3233/JIFS-234709

Wu, K., Li, L., Li, J., & Li, T. (2013). Ontology-enriched multi-document summarization in disaster management using submodular function. *Information Sciences*, *224*, 118–129. https://doi.org/10.1016/j.ins.2012.10.019

Xiang, J., Liu, M., Li, Q., Qiu, C., & Hu, H. (2024). A cross-guidance cross-lingual model on generated parallel corpus for classical Chinese machine reading comprehension. *Information Processing & Management*, *61*(2), 103607. https://doi.org/10.1016/j.ipm.2023.103607

Xie, T., Zhen, Y., Li, T., Li, C., & Ge, Y. (2021). Self-supervised extractive text summarization for biomedical literatures. In 2021 IEEE 9th international conference on healthcare informatics (ICHI) (pp. 503–504). https://doi.org/10.1109/ICHI52183.2021.00091

Xiong, J., Yang, J., Yan, L., Awais, M., Khan, A. A., Alizadehsani, R., & Acharya, U. R. (2024). Efficient reinforcement learning-based method for plagiarism detection boosted by a population-based algorithm for pretraining weights. *Expert Systems with Applications*, *238*, 122088. https://doi.org/10.1016/j.eswa.2023.122088

Yadav, A. K., Ranvijay, S., Yadav, R. S., & Maurya, A. K. (2023). State-of-the-art approach to extractive text summarization: A comprehensive review. *Multimedia Tools and Applications*, *82*, 29135–29197. https://doi.org/10.1007/s11042-023-14613-9

Yang, C., Wang, C., Zheng, R., & Geng, S. (2023). Link prediction in research collaboration: A multi-network representation learning framework with joint training. *Multimedia Tools and Applications*, *82*, 47215–47233. https://doi.org/10.1007/s11042-023-15720-3

Yang, M., Tu, W., Qu, Q., Lei, K., Chen, X., Zhu, J., & Shen, Y. (2019). MARES: Multitask learning algorithm for web-scale real-time event summarization. *World Wide Web*, *22*, 499–515. https://doi.org/10.1007/s11280-018-0597-7

Yang, S., Yim, J., Kim, J., & Shin, H. V. (2022). CatchLive: Real-time summarization of live streams with stream content and interaction data. In Proceedings of the 2022 CHI conference on human factors in computing systems (pp. 1–20). https://doi.org/10.1145/3491102.3517461

Yang, Y., Zhang, C. X., & Rickly, J. M. (2021). A review of early COVID-19 research in tourism: Launching the annals of tourism research's curated collection on coronavirus and tourism. *Annals of Tourism Research*, *91*, 103313. https://doi.org/10.1016/j.annals.2021.103313

Yang, Z. G., Agócs, Á., Kusper, G., & Váradi, T. (2021). Abstractive text summarization for Hungarian. *Annales Mathematicae et Informaticae*, *53*, 299–316. https://doi.org/10.33039/ami.2021.04.002

Yao, K., Zhang, L., Luo, T., & Wu, Y. (2018). Deep reinforcement learning for extractive document summarization. *Neurocomputing*, *284*, 52–62. https://doi.org/10.1016/j.neucom.2018.01.020

Yip, W. S., Yan, H. E., Zhang, B., & To, S. (2024). The state-of-art review of ultra-precision machining using text mining: Identification of main themes and recommendations for the future direction. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *14*(1), e1517. https://doi.org/10.1002/widm.1517

Yoshida, Y., Suzuki, J., Hirao, T., & Nagata, M. (2014). Dependency-based discourse parser for single-document summarization. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP) (pp. 1834–1839). https://doi.org/10.3115/v1/D14-1196

Yuan, C., Bao, Z., Sanderson, M., & Tang, Y. (2020). Incorporating word attention with convolutional neural networks for abstractive summarization. *World Wide Web*, *23*, 267–287. https://doi.org/10.1007/s11280-019-00709-6

Zaman, F., Shardlow, M., Hassan, S.-U., Aljohani, N. R., & Nawaz, R. (2020). HTSS: A novel hybrid text summarisation and simplification architecture. *Information Processing & Management*, *57*(6), 102351. https://doi.org/10.1016/j.ipm.2020.102351

Zhang, H., Song, H., Li, S., Zhou, M., & Song, D. (2023). A survey of controllable text generation using transformer-based pre-trained language models. *ACM Computing Surveys*, *56*(3), 1–37. https://doi.org/10.1145/3617680

Zhang, M., Lu, J., Yang, J., Zhou, J., Wan, M., & Zhang, X. (2024). From coarse to fine: Enhancing multi-document summarization with multi-granularity relationship-based extractor. *Information Processing & Management*, *61*(3), 103696. https://doi.org/10.1016/j.ipm.2024.103696

Zhang, R., Li, W., & Gao, D. (2013). Towards content-level coherence with aspect-guided summarization. *ACM Transactions on Speech and Language Processing*, *10*(1), 1–22. https://doi.org/10.1145/2442076.2442078

Zhang, T., Ladhak, F., Durmus, E., Liang, P., McKeown, K., & Hashimoto, T. B. (2024). Benchmarking large language models for news summarization. *Transactions of the Association for Computational Linguistics*, *12*, 39–57. https://doi.org/10.48550/arXiv.2301.13848

Zhang, X., Wei, F., & Zhou, M. (2019). HIBERT: Document level pre-training of hierarchical bidirectional transformers for document summarization. In Proceedings of the 57th annual meeting of the association for computational linguistics (pp. 5059–5069). https://doi.org/10.18653/v1/P19-1499

Zhao, B., & Xing, E. P. (2014). Quasi real-time summarization for consumer videos. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2513–2520). https://doi.org/10.1109/CVPR.2014.322

Zhao, Z., Alzubaidi, L., Zhang, J., Duan, Y., & Gu, Y. (2023). A comparison review of transfer learning and self-supervised learning: Definitions, applications, advantages and limitations. *Expert Systems with Applications*, *242*, 122807. https://doi.org/10.1016/j.eswa.2023.122807

Zheng, L., Shen, C., Tang, L., Li, T., Luis, S., Chen, S. C., & Hristidis, V. (2010). Using data mining techniques to address critical information exchange needs in disaster affected public-private networks. In Proceedings of the 16th ACM SIGKDD international conference on knowledge discovery and data mining (pp. 125–134). https://doi.org/10.1145/1835804.1835823

Zheng, S., Li, Z., Wang, J., Qu, J., Liu, A., Zhao, L., & Chen, Z. (2023). Long-document cross-lingual summarization. In Proceedings of the sixteenth ACM international conference on web search and data mining (pp. 1084–1092). https://doi.org/10.1145/3539597.3570479

Zhu, C., Yang, Z., Gmyr, R., Zeng, M., & Huang, X. (2021). Leveraging lead bias for zero-shot abstractive news summarization. In Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval (pp. 1462–1471). https://doi.org/10.1145/3404835.3462846

How to cite this article: Chen, X., Xie, H., Tao, X., Xu, L., Wang, J., Dai, H.-N., & Wang, F. L. (2024). A topic modeling-based bibliometric exploration of automatic summarization research. *WIREs Data Mining and Knowledge Discovery*, *14*(5), e1540. https://doi.org/10.1002/widm.1540