Complex Systems Informatics and Modeling Quarterly (CSIMQ)
eISSN: 2255-9922
Published online by RTU Press, https://csimq-journals.rtu.lv
Article 216, Issue 39, June/July 2024, Pages 86–104
https://doi.org/10.7250/csimq.2024-39.05

Survey on Organizational Chat Conversation Analysis: Exploring Dialogue Summarization from a Knowledge Discovery Perspective

Ksenija Kosilova* and Ilze Birzniece

Institute of Applied Computer Systems, Riga Technical University,
6A Kipsalas Street, Riga, LV-1048, Latvia

Ksenija.Kosilova@edu.rtu.lv, Ilze.Birzniece@rtu.lv

Abstract. With the latest advances in natural language processing technologies, multi-participant conversation summarization features are now embedded in the most widely used collaboration platforms offered by industry leaders such as Microsoft, Google, and Zoom. This allows employees to streamline their work and increase efficiency by summarizing long chat threads. In this study, an attempt has been made to perceive summarized chat conversations as a tool for knowledge discovery and reusable information extraction within an organization in general or during projects. To this end, recent scientific articles have been reviewed to identify the most effective techniques and approaches for summarizing chat threads and conversations alongside the challenges and peculiarities of collaborative text-based communication. In addition, significant attention has been paid to the further utilization of the extracted information to represent the knowledge for further reuse.

Keywords: Dialogue Summarization, Collaborative Chat Conversation Summarization, Knowledge Extraction, Information Extraction, Reusable Knowledge.

1 Introduction

The motivation of organizations to utilize text summarization tasks for condensing multiparticipant chat conversations is to assist employees with accomplishing imminent but repetitive tasks, reducing time spent on low-value-added activities such as searching for useful information in long chat threads, and consequently increasing employees' work pace. Driven by the valuable content and the typically unstructured and fragmented nature of chat conversations within teams and organizations, the potential for more effective reuse of extracted information and knowledge must be explored.

In this research, we aimed to provide insights into the challenges and limitations of methods and techniques applicable to reusable knowledge extraction from chat conversations. In addition,

* Corresponding author

Reference: K. Kosilova and I. Birzniece, "Survey on Organizational Chat Conversation Analysis: Exploring Dialogue Summarization from a Knowledge Discovery Perspective," Complex Systems Informatics and Modeling Quarterly, CSIMQ, no. 39, pp. 86–104, 2024. Available: https://doi.org/10.7250/csimq.2024-39.05

Additional information. Author's ORCID iD: I. Birzniece – https://orcid.org/0000-0002-5775-6138. PII S225599222400216X. Article received: 27 March 2024. Accepted: 25 June 2024. Available online: 31 July 2024.

knowledge discovery and reusable knowledge extraction from a variety of communication mediums have been examined from the perspective of organizational knowledge sources. The analysis of unstructured chat conversations, stored in team and public chat mediums, is driven by the vast amount of reusable knowledge, its structural complexity, and, most importantly, the potential for more effective reuse of extracted information [1]. Various researchers argue that chat conversations as reusable knowledge can be used as a source of information [2], personal experiences of developers, and software artifacts such as code or documentation [3]. For instance, in the area of telemedicine, medical conversation summarization mainly addresses the issue of repetitive conversations among doctors. An equally significant aspect of the knowledge perspective of conversation summarizations is their representation in knowledge-intensive industry domains such as software development and telemedicine, where they have been clearly represented as ontologies and commonsense knowledge, suggesting that domain-specific concepts and terms are crucial for quality information extraction from chat conversations and subsequent reuse for productive workflows.

It is worth mentioning that significant scientific community efforts have been made to comprehend the role of instant messaging (IM) in knowledge management (KM) and its potential to enhance organizational productivity and facilitate the learning process. In this article, the literature survey seeks to identify the relationship between reusable knowledge, commonsense knowledge, knowledge discovery, artifacts, and sources. Knowledge artifacts (KA) are concentrated forms of industry or project-specific knowledge. Transforming extracted useful information into knowledge through knowledge discovery and sharing is suspected to be iterative.

Nevertheless, summarizing multiparty conversations with interactive communication flow can be challenging due to multiple speakers and the complex flow of conversational content [4]. Multiple speakers' asynchronous chats typically cover various topics that may continuously drift. Moreover, multiple threads of conversation in a single sequence of messages might occur [5].

A review of related scientific literature shows that the medical and software development industries have the highest demand for quality summarization. However, they face specific challenges finding uniform solutions due to complex discussions and informal language and variety of conversation platforms.

In this article, the survey was conducted based on the approach proposed in [6], which supported the review process of 16 shortlisted articles and provided the following contributions: (1) overview of summarization utilization as a useful source from a knowledge perspective and at the same time as an input for KA creation for further information extraction; (2) overview of the latest developments in multiparty chat conversation summarization, highlighting existing challenges and objectives for future work.

The article is structured as follows: Section 2 defines the goal and research questions and describes the survey methodology employed. Section 3 is dedicated to the research of the knowledge perspective of dialogue summarization and describes its outcomes. Section 4 is dedicated to the research of challenges and methods utilized in dialogue summarization. Section 5 concludes the article and presents the conclusions drawn from the preceding two sections and the prospects for future work.

2 Survey Methodology

A systematic literature review facilitates meaningful scientific literature research by disentangling accumulated knowledge with appropriate breadth, depth, thoroughness, and consistency [6]. It contributes to the practical synthesis and analysis of how previous related research builds on each other. It establishes a strong foundation of knowledge for developing assumptions, identifying and addressing gaps in research domains, and pointing out the areas that need further investigation [7].

The goal of the survey is to review the prospects of text-based multiparticipant conversation summarization and chat text analytics applicable to formal and informal communication in organizations. It aims to provide insights into the challenges and limitations of methods and techniques applicable for extracting useful information and chat conversation summarization. Moreover, useful information extraction and knowledge discovery from a variety of text-based communication mediums must be examined from the perspective of organizational knowledge sources. To achieve the goal of this survey, the following general research questions have been formulated:

RQ1. How knowledge discovered from analyzed and summarized chat conversations between employees or business actors can be beneficial, efficiently and cost wise reasonable to extract?

RQ2. What are the most suitable methods and techniques for chat analysis and summarization tasks applicable to formal and informal text-based communication?

RQ3. What are the challenges, obstacles, most noteworthy advancements, and trends of text summarization methods of multiparticipant chat analysis on extracting reusable organizational knowledge?

Based on the goal of the survey and research questions formulated, the following keyword combinations representing notions and terms related to the current survey were used for reference during the scientific literature search:

1. ("Chat summarization" OR "summarization") AND "information extraction"; ("chat summarization" OR "dialogue summarization" OR "summarization") AND ("knowledge extraction" AND "information extraction"); ("information extraction" OR "text summarization") AND ("text-based communication" OR "dialogue").
2. "Conversation summarization" OR "dialogue summarization"; "summarization" AND ("knowledge extraction" OR "knowledge management" OR "knowledge artifact") OR "help desk chat".
3. "Challenges" AND "dialogue summarization".
4. Other variations of notions and terms relevant in the context of the research questions of the survey.

Backward reference search, backward authors search, and previously used keyword search techniques were employed successfully and helped to follow methods, techniques, challenges, and, most significantly – research streams [6]. Due to the information system literature's multidisciplinary and diverse nature, it was challenging to exhaust literature resources on the survey topic. Therefore, Scopus and ScienceDirect were used by employing a primary keyword search approach while refining keyword combinations and extracting new ones to avoid using only buzzwords, thus narrowing the scientific foundation of the survey.

To avoid a narrow literature background and shallow depth, a secondary keyword search was applied in the following high-quality scientific indexes: IEEE Xplore Digital Library, ACM Digital Library, SpringerLink, and ACL Anthology. Meaningful keyword search technique has been utilized in the first place for preliminary article selection while limiting search output by literature type to conference papers, studies, research, and articles, and by subject – to Computer Science, Decision Sciences, Knowledge Systems, Computational Linguistics. The articles published earlier than 2015 have been filtered out due to the rapid development of methods in this field, making older methods obsolete. The following inclusion criteria were put in place for adding an article to the general list of surveyed articles for further refinement to be used as a foundation literature:

1. Peer-reviewed articles in English.
2. The articles provide conclusions usable for practitioners, define future work, or both.
3. At least one of the following topics should be covered to include the article in the research scope of the survey:
   - The article considers organizational text-based communication as a source of reusable knowledge or presents the solution for useful information extraction;
   - The article considers text summarization methods applicable to IM applications, collaborative platforms, group chats, and personal chats, preferably in the context of organizations or teams;

- The article considers employees' or chat participants' mutual connections;
- The article discusses trends in conquering challenges and obstacles related to the application of summarization methods.

To facilitate the inclusion of high-quality articles and to administer their relevance to established research questions, the following exclusion criteria were put in place for article inception into the principal list of surveyed articles:

1. The article did not provide a substantial response to any of the research questions, or its overall quality was not satisfactory.
2. The text-based conversation summarization and organizational chat analysis were not sufficiently presented or addressed: social media text summarization, lengthy text documents summarization, and meeting transcript summarization.
3. Large-scale surveys on text-based conversation summarization and chat analysis methods.

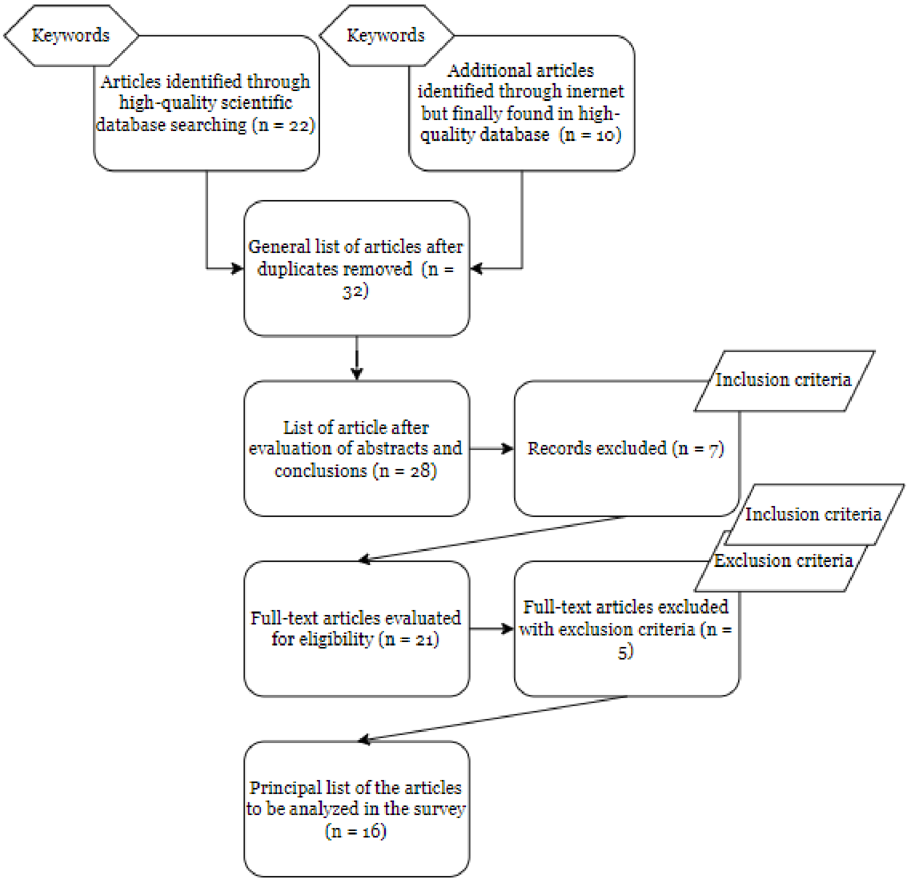

Figure 1. The Flow diagram representing the article selection process

Primarily, 32 articles were selected to be included in the general list, as seen from the algorithm schema depicted in Figure 1. The abstract and conclusions of each article have been read and evaluated per the defined inclusion and exclusion criteria. Hence, 21 articles were shortlisted and read thoroughly, and finally, 16 were selected for inclusion in the principal survey list.

The survey scope was not limited by industry or domain since the primary purpose of it was to explore as many diverse chat and discussion structures as possible with multiple or two participants. Citation ranking was not considered as some of the articles were released recently and excluding them based on citation ranking would result in an insufficient sample size for the survey. The shortlisted articles were published between 2017 and 2023 in the following journals, conferences, proceedings, and workshops: The Journal of Systems & Software, Applied Sciences Journal, International Conference on Biomedical and Health Informatics, International Conference on Mining Software Repositories, International Conference on Research and

Development in Information Retrieval, ACM on Human-Computer Interaction, International Conference on Bioinformatics and Intelligent Computing, International Conference on Big Data, Artificial Intelligence and Risk Management, Annual Conference of the International Speech Communication Association, International Joint Conference on Neural Networks, CEUR Workshop, International Conference on Computing and Network Communications, International Conference on Algorithms, International Workshop on Multimedia-based Educational and Knowledge Technologies for Personalized and Social Online Training.

The principal list of surveyed articles consisted solely of the ones selected using inclusion and exclusion criteria. However, articles excluded from the principal list have still been used for terminology references or argumentation in the survey pursuant to its goal. In addition, it was necessary to reference literature on KM to define terms and concepts within the body of knowledge properly. However, even a general list of articles did not cover both aspects – KM and text summarization.

The conducted survey was focused on the summarization of text and information extraction from IM applications, discussion threads, and chat rooms in collaborative discussion platforms such as Gitter[†], Slack[‡], and GitHub Discussions[§], conversations from issue tracking systems, patient-doctor conversations in the context of organizations, teams, or individual text-based communication.

3 Knowledge Perspective of Chat Analysis in Organizations

According to Microsoft's "Work Trend Index Annual Report" [8] released in May 2023, the intensity of communication, diversity, and volume of information within past years have increased fiercely, resulting in employees' struggles to keep up with the pace and information overload. It has been revealed that 68% of respondents lack uninterrupted focus during work, while 62% spend too much time searching for necessary information [8]. The report [8] also highlights the concept of perceiving meeting summaries, chats, and emails as digital artifacts.

Major collaboration platforms, including Zoom, Google Spaces, and MS Teams, propose summarization as a cutting-edge solution to cope with text-based communication overload. This aids in managing everything stored and shared on the platforms' clouds: meetings, emails, whiteboards, and more.

In November 2023, Microsoft released the Teams chat conversation summarization feature as a part of Microsoft 365 Copilot global functionality to help users deal with information flow across all Teams chat threads, aided by AI-powered automatically generated summaries of conversations. Moreover, Copilot is truly ubiquitous owing to the usage of Microsoft Graph (users' emails, calendar, chat threads, documents, meetings, conversations) data analyzed by Semantic index, which uses Large Language Models (LLMs) for content and relationships mapping in addition to data from the web. Copilot uses pre-trained LLMs such as Generative Pre-Trained Transformers (GPT) and extensive datasets [9].

Earlier in October 2022, Google released the Chat Conversation Summaries feature for Google Workspace business accounts. Once on-demand summaries are available, posters containing various topics discussed are automatically generated. The feature uses Pegasus, a transformer abstractive summarization model employing knowledge distillation. The developers had to resolve quality issues in the summaries, such as misattribution of utterances between participants and misrepresentation of the conversation, due to user problem reports [10].

In September 2023, Zoom AI Companion was launched, offering a Team Chat Thread Summary tool to condense chat discussions and conversation threads. Zoom uses the content of the conversation and participant names to create summaries by leveraging third-party models and

[†] https://gitter.im/
[‡] https://slack.com/
[§] https://github.com/features/discussions

proprietary distinctive methods by dynamically integrating intrinsic LLM with Meta Llama 2, OpenAI, and Anthropic [11].

3.1 Knowledge Discovery and Extraction Aspects Within the Organizational Environment

In recent decades, researchers have invested significant efforts in delving into KM and its underlying theory, which entails handling processes of transforming data into information and eventually into knowledge. This includes creating and acquiring, transferring, and sharing knowledge within the context of organizational operations and performance improvements. The impact of information technology advancements on knowledge management in organizations has been significant. This has resulted in the creation of various tools such as KM systems, knowledge-sharing platforms, enterprise group chat collaboration tools, and integrated business intelligence tools. Artificial intelligence is also being used in KM strategies and practices, automating knowledge processes, extracting valuable insights, and facilitating informed decision-making. Furthermore, preceding scientific literature on KM, among other directions, is focused on developing organizational culture and employee engagement in knowledge-sharing environments, forming a reliable basis for further research endeavors, and establishing a relationship between sharing knowledge, information extraction, and collaborative messaging.

According t o [12], information becomes knowledge when integrated into a network of semantic connections. This knowledge can then be applied to comprehend a situation or work towards achieving a goal, as desired by an organization. The definition of knowledge emphasizes its close relationship with emerging technologies. This implies that the latest developments and tools should be used to optimize the organization's internal knowledge resources.

From the perspective of data analytics, knowledge discovery is the process of discovering useful knowledge in a broad range of sources, such as relational databases, images, or texts [2]. According to [13], knowledge discovery is one of the phases in the context of KM practices delineating the transformation of raw data from various sources into fundamental information, signifying novel explicit knowledge by utilizing preprocessing techniques, computational algorithms, and statistical models. Once knowledge has been captured, it must be represented in some form, for instance – ontologies or folksonomies [9]. Folksonomies are structures of knowledge that consist of tags created by users and content resources. In contrast to taxonomies, which are hierarchical and established not by the user but by an authority, folksonomies are more flexible and typically embedded in local cultural and social systems.

From a KM perspective, a KA is an object that holds a representation of knowledge, such as documents or files [14], ontologies, folksonomies, or other constructs, and is often designed to separate knowledge from its use. Managing artifacts can potentially add value to the organization by improving accessibility to its knowledge resources. From this perspective, KAs serve as a means for sharing knowledge. Salazar-Torres et al. [14] proposed approaching the notion of a KA from an artificial intelligence perspective, defining it as an artifact primarily composed of knowledge. KA are artifacts that represent knowledge by encoding it and require knowledge input for its creation, being the incarnation of knowledge itself. Commonsense knowledge refers to the ability of a computer system to understand and process information that is considered to be common sense.

Notable examples of KA include ConceptNet, a large-scale commonsense knowledge base employed in the development of summarization models by Tiwari et al. [15] and Xiachong et al. [16], and the Retrieval-Augmented Generation (RAG) framework, which was not included in the general list of articles but is a valuable method to mention for knowledge extraction from conversations. RAG is a framework that exploits the benefits of conventional information retrieval systems, such as databases, while also capitalizing on capabilities of LLMs. RAG employs search algorithms to request external data (e.g., websites, databases, knowledge bases) in order to obtain pertinent information and perform pre-processing [17]. Subsequently, the previously obtained information is incorporated into the LLM, thereby augmenting the context and enabling the LLM to generate responses that are more precise and accurate. The RAG models, comprising the RAG-sequence and RAG-token models, developed by a group of researchers in 2020 [18] and presented by Lewis et al. [18], combined the parametric memory contained in a pre-trained seq2seq model with the non-parametric memory accessed by retrieval systems. In contrast, traditional LLMs are constrained by their pre-trained knowledge and data sets [17]. Ahn et al. [19] put forth an enhanced methodology for knowledge-based response generation, one that is grounded in external knowledge and the contextual nuances of the conversational setting. This approach entailed the retrieval of an appropriate set of documents that were relevant to both the topic and the local context of a conversation and employed a novel data weighting scheme to prompt the model to generate knowledge-based responses in the absence of ground truth knowledge. A dataset of conversation threads extracted from Reddit.com was used to test the methodology.

In this way, RAG may assist LLMs with access to curated organizational knowledge bases, thereby ensuring that generated summaries of project or work-related conversations are grounded in factual information. However, if the underlying data sets lack consistent organization, categorization, and metadata, retrieval algorithms will be unable to identify the most valuable information. It is, therefore, critical for conversation summarization to ensure that data is properly structured, tagged, and accessible, posing additional challenges for the summarization of long threads of conversations. The primary limitation of RAG models is their inability to fully comprehend whether the retrieved data is the most pertinent information required by the language model to resolve the query effectively. RAG models are still effective for summarizing large text documents. This involves the retrieval of relevant information from a variety of sources in order to produce articles, reports, and concise summaries of the highest quality. Additionally, RAG models facilitate the enhancement of conversational agents, such as customer service chatbots, virtual assistants, and other conversational interfaces.

The interaction between organizational knowledge and data exchange processes has been investigated thoroughly, aided by information categorization and subsequent conceptual modeling of organizational business processes. In this context, corporate memory has been proclaimed an information flow type and was significant for facilitating organizational knowledge about products [20], stemming from operational activity and business process analysis. In essence, corporate memory is involved in sharing and transferring knowledge within the organization regarding its products and services, thereby optimizing internal data exchange processes.

3.2 Text-based Communication as a Source of Organizational Knowledge

Organizations can acquire knowledge from their employees, customers, suppliers, and computer systems – participant's knowledge. This knowledge is equipped with manipulation capabilities [21], [22], unlike KAs embedded in electronic or analog media, such as documents, which lack these capabilities. One of the relevant aspects of the sourcing of organizational knowledge is its representation through products or services provided by the organization, as they should not only be considered as the result of capital, material, and labor inputs and efforts or, as the authors of [21] put it: "products in an organization's inventory are artifacts that represent the knowledge used to build them." Organizational knowledge is also influenced by its environment, which includes government, media, rival entities, and other institutions.

Later, practitioners began to consider the peculiarities of adopting IM in the workplace and its role in creating, sharing, and retaining knowledge for subsequent reuse. Researchers began to apply techniques to extract knowledge from IM [1] and professional emails [23] produced in the course of project realization to structure concepts during project implementation, dealing with such a variety of knowledge as project memory, exemplifying the organizational and collaborative dimensions of knowledge. The notion of reusable knowledge refers to information that can be applied in various contexts or for solving similar problems and for future projects.

Recent research in software engineering and maintenance has acknowledged the value of information stored within discussion threads. It has been proposed to extract knowledge from Q&A

forums such as Stack Overflow [24] and reusable software development knowledge contained in IM [3]. According to [25], software maintainers move discussions manually from live chat on Discord to GitHub Discussions to preserve valuable information and insights for improving the project later. Based on the above, it can be concluded that text-based communication must be perceived as a source of information within a project or organization and can be transformed into knowledge represented by KA using knowledge discovery techniques.

3.3 Significance of Text-based Communication in Knowledge Discovery, Transfer, Sharing, and Reuse

Apart from chat conversation, knowledge reuse for performance optimization within teams and organizations during project development is argued by Wanderley et al. [2]. They claim that knowledge delineated by folksonomies can be used to interpret new utterances in a dialogue and consequently utilized for analysis, trends detection, conversations categorization, group identification, behavior prediction, and another insight gain.

By mining knowledge from Q&A forums, the software engineering community intends to support integrated development environments (IDE) recommendation learning and API recommendations, automatically generate comments for source codes, and build thesauri and knowledge graphs of software-specific terms and terms commonly used in software engineering [24]. For instance, Slack channel transcripts can be studied to identify trending discussion topics in a programming community and understand common challenges and misconceptions among developers [24]. As a result, these studies would be guidance for future research, such as developing software support and maintenance tools.

Taking care of IM and live chat content, software maintainers prevent the loss of important knowledge that may occur in the transient nature of live chat. This also ensures that absent developers can benefit from the discussions, promoting knowledge retention within the software collaborative platform, considering alternatives, and discussing problems [25]. Moreover, aspects of decision-making, nourishing project documentation, and software development knowledge reuse are essential [25]. Developers seek knowledge from IM in real time to obtain feedback from experts who can share their expertise with them operatively.

Alternatively, taking into consideration the instant issue commenting feature in the working environment of Issue Tracking Systems (ITS) allows users from different backgrounds to create discussions around bug fixes and software improvements. Gilmer et al. [26] stress that issue discussion threads should be perceived as dynamic project documentation that records abundant information about collaborative progress.

In the field of telemedicine, patients can have private chats with clinicians to obtain medical advice and avoid long waiting times. Patients' medical cases are stored efficiently, enhancing their reusability during further treatments and reducing the cognitive load of clinicians [15], [27]. Automating routine activities in the medical domain, such as generating diagnosis reports and preliminary symptom investigation reports, addresses the issue of repetitive conversations among clinicians.

3.4 Knowledge Perspective of Chat Conversation Analytics and Summarization

It has been revealed that notions of KA, knowledge discovery, knowledge resources, reusable knowledge, commonsense knowledge, knowledge sharing, and creation are interrelated. Reusable knowledge and knowledge artifacts are concentrated forms of industry or project-specific knowledge. Their cyclical nature is characterized by improvements in extracted information, which becomes useful after applying knowledge resources and KA, which can be utilized for knowledge infusion [15], [16] in the summarization task. Transforming extracted useful information into reusable knowledge through knowledge discovery and sharing knowledge are iterative. This is because the knowledge is directed to improve processes and enhance information accessibility for reuse in text-based communication in the context of collaborative environments. Figure 2 below distills the knowledge perspective of chat analysis in organizations from previously discussed research and depicts the iterative process of knowledge reuse in the context of a chat conversation summarization task.

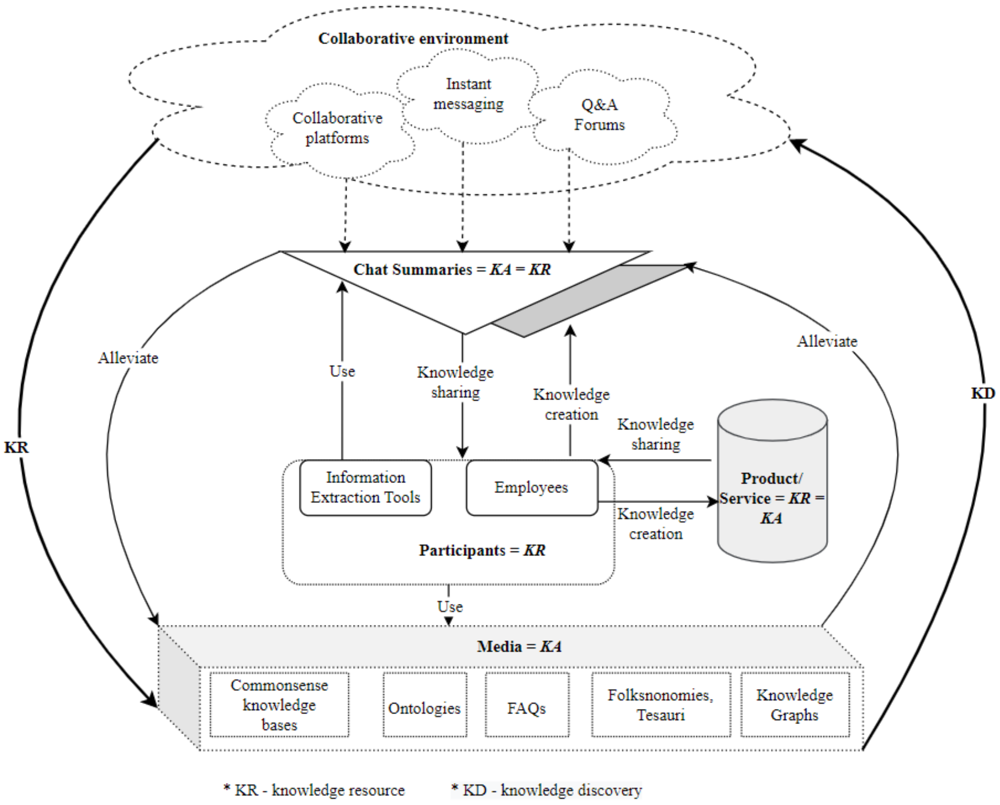

Figure 2. Iterative circle of information extraction, reusable knowledge creation, sharing, and discovery

An equally significant aspect of the knowledge perspective is its representation in knowledge-intensive industry domains such as software development and telemedicine. As evidence, they were apparently represented as ontologies and commonsense knowledge, suggesting that domain-specific concepts and terms are crucial for quality information extraction from chat conversations and subsequent reuse for productive workflows. For instance, information delineated by folksonomies created using text-based communication analysis can be used to interpret new utterances in collaborative IM platforms and consequently utilized for further analysis, software development trend detection, categorization, software developer behavior prediction, and other insights. Thematic mapping of IM content to the Software Engineering Body of Knowledge (SWEBOK) could create a more sophisticated thematic classification for improving dialog summarization tools, thus creating more high-quality reusable knowledge for developers.

The products and services created in a collaborative environment serve not only as deliverables but also as a source of reusable knowledge for their subsequent maintenance and enhancing project implementation practices. Employees, in turn, contribute to creating KAs using their acquired knowledge, establishing a cyclical process that increases knowledge through interaction with these artifacts. Summarized chats play a key role in reinforcing this iterative cycle, contributing to the continuous improvement of knowledge in a collaborative context. Table 1 shows a condensed summary of knowledge-related concepts discovered in the surveyed articles, including the source of information, aspects of observing reusable knowledge, and the knowledge discovery aspects represented by the utilized and created KA.

From articles surveyed and displayed in Table 1, 10 articles consider chat content and its summarization as a source of valuable information and, after significant processing, as reusable knowledge. However, the rest of the articles surveyed consider summaries as a tool for combating information overload and avoiding repetitive tasks, but at the same time, they emphasize commonsense knowledge infusion for the summarization task itself. The third, fourth, and fifth columns of the table represent the source of organizational data input, knowledge artifacts utilized for chat analysis and/or summarization, and KA in the form of thesauri, FAQs, instructions, and specific knowledge graphs.

Table 1. Overview of reusable knowledge aspects, sources of information, and knowledge artifacts used and created during chat analysis

Surveyed articles	Reusable knowledge aspect discussed	Source of organizational information	Knowledge artifact utilized	Knowledge artifacts created or prototypes created
[3]	✓	Discussions on Gitter live chat rooms and GitHub Discussions	Thematic mapping of IM content to SWEBOK	
[4]		Not specified		
[5]		WhatsApp multiparty asynchronous chats		
[15]	✓	Clinician-patient conversations	ConceptNet for external knowledge infusion into the summarization model	Annotations include medical entities, medical departments, disease tags
[16]	✓	Not specified	ConceptNet to cope with utterances mapping	
[24]	✓	Conversation in public Slack channels		Thesauri, software-specific terms knowledge graphs, IM themes mapping for API documentation
[25]	✓	Gitter live chat rooms and GitHub Discussions	Knowledge encoded in the pre-trained BERT model	
[26]	✓	Software development discussions on ITS		
[27]	✓	Not specified		
[28]				
[29]	✓	Not specified	Stanford CoreNLP’s Open Information Extraction tool – OpenIE	
[30]				
[31]	✓	Group chats at collaborative platforms		
[32]				
[33]			Knowledge encoded in the pre-trained BERT model	
[34]	✓	Common project coordination chats		How-To Instructions, FAQs, Summary reports

4 Chat Analytics and Dialogue Summarization

The application of machine learning algorithms and behavioral analysis is a common approach to exploratory analysis of long organizational chat threads. This analysis can extract insights such as the most active days, the top active users, included URLs and media content, employee workload and other chat participation metrics. Researchers have also been concerned with the development of a map of themes that could be applied in approaches for augmenting software documentation or semantic tagging of chat rooms. Chat summarization is normally used for labeling purposes and, in some cases, to eliminate the need for manual annotations of extensive thread conversations.

4.1 Extractive and Abstractive Summarization

Text summarization is a common task in natural language processing that seeks to condense a source text into a summary encapsulating the main idea. Text summarization applies to literature, news outlets, media, legal documents, scientific papers, etc. Research on text summarization techniques has been ongoing for a significant period, with two main categories based on the generated text source: extractive and abstractive methods. Extractive methods choose word blocks from the source text to create summaries. In contrast, abstractive methods generate summaries more prolifically, incorporating words outside the source text [28] and using out-of-vocabulary words. Extractive methods usually produce detailed but less general summaries, opposing abstractive methods, which are more general but potentially more error-prone.

Another key fact is the variety of summarization task approaches and their peculiarities. Namely, feature-based summarization approaches extract sentence and word-level features and formulate summarization as a trainable model. Query-based summarization approaches select sentences semantically related to one or more user-provided query sentences. These models use similarity measuring methods such as the vector space model or sentence embedding to retrieve and summarize a document that satisfies the input information invoked by a user's query [29]. In a nutshell, query-based conversation summarization is based on phrasal queries reflecting user-required information. In contrast, graph-based methods extend the TextRank algorithm to model text as a graph, where sentences are represented as nodes and words overlap as edges.

4.2 Specifics of Dialogue and Chat Conversation Summarization

Firstly, dialogue and chat conversations are forms of dynamic information exchange, and their main characteristics relate to the most ferocious summarization task challenges. Dialogue and chat conversations can be informal [24], [27], [28], [29], iterative [4], asynchronous [24], changing topic of conversation, lack of paragraph separation [30], [31], and interspersed with backchanneling [27], [32], reaffirmations, and speaker interruptions [24], [31], [32]. Additionally, utterances may come from different interlocutors [4], [28], [32], leading to topic drifts and lower information density.

Unlike the common summarization tasks on generic texts, dialogue and chat conversation summarization is deemed much more challenging since general domain text techniques are out of multiple speakers modeling facet [28] because a sole person typically writes general texts, while conversations involve multiple speakers interacting. Moreover, there are unique characteristics inherent to the dialogue text writing style. Generic texts tend to be formal, while dialogue texts often include casual language, colloquial expressions, and code-mixed language. Chat conversations tend to have unusual features, such as abbreviations, acronyms, emoticons, and misspelled words [27]. In addition, the repetition of names of interlocutors and necessary actions significantly decreases the quality of summarizations [32].

One of the specific problems mentioned by Chatterjee et al. [24] concerning information in Q&A forums was the unstructured nature and absence of predefined delineation of conversations knowledge sharing manner. In addition to the previously mentioned characteristics, there are multiple conversation threads in a single sequence of chat messages [5]. Therefore, as Sinha et al. [5] stated, such aspects as discussion threads, time windows, and topic/sub-topics must be considered for summarization. As a result, the scope of summarization can vary. It can be over a single thread, over a duration of time, or over a given number of messages that encompass a duration of time, as well as multiple threads and topics.

Nonetheless, the dialogue summarization task has progressed in several directions, including feature-based extractive summarization, recurrent neural network-based summary generation, and pre-trained large language model-based summarization. In recent years, the focus has been on aspect-guided dialog summarization, namely, domain, intent, and keyword [15].

4.3 Dialogue and Chat Conversation Summarization Challenges Addressed

Besides previously mentioned dialogue summarization peculiarities and challenges, Li et al. [29] address issues related to factual inaccuracy/inconsistency of summarizations. The authors argue that summarization methods usually produce summaries with high matching metrics, such as ROUGE, but tend to be inaccurate from the content perspective.

In addition, summarizing chats within the medical advising domain imposes additional challenges, such as requiring knowledge of domain-specific terms. According to [27], infusing knowledge of domain-specific terms helps to handle a higher frequency of misspelled medical terms and assess the medical importance of information so that critical information is not missed. Tiwari et al. [15] address the critical significance of visuals in telemedicine.

Within the software development domain, especially open-source software projects relying heavily on asynchronous remote collaboration, manual summarization is a common strategy for developers to contribute to discussions within long-living issue threads. However, summaries often get drowned under subsequent comments and fail to reach the target readers. Therefore, Gilmer et al. [26] address the lack of key design and feature requirements for summarization applications to bridge the user interface gap. Silva et al. [3] successfully attempted to analyze the relevance of the knowledge accumulated in developers' chat rooms and emphasized the necessity to understand primarily the themes discussed in chat rooms.

The increase in scale and computational overhead of large pre-trained models has led to the demand for compressing these models into smaller versions. Zhao et al. [32] challenged the objective of maintaining high precision while achieving faster execution time for inference.

4.4 Chat Analytics and Dialogue Summarization Techniques and Significant Outputs

To address factual inaccuracy, Li et al. [29] proposed a fact-augmentation mechanism to infuse factual information from chat conversations. The model incorporated the fact graph derived from the dialogue itself into the summarization generation process and enhanced the gain of factual information through the fact-sensitive scoring element [29]. The authors effectively utilized the ground fact to resolve the factual mismatch between the conversation and the generated summarization.

In medical conversations between patients and doctors, [27] proposed aiding the open-source natural language processing system, cTAKES, which extracts clinical information from unstructured text, such as dialogue, using medical concepts from the UMLS ontology. Authors applied a feature-based method and achieved more accurate summarizations with external knowledge infusion, while Tiwari et al. [15] proved that their knowledge (ConceptNet) augmented model enhanced with context-aware modality-driven fusion method demonstrated a remarkable advantage in labeling clinical conversations with medical departments and diseases for further diagnosis. Tiwari et al. [15] asserted that multi-modal summarization achieved coherent and vital information from multiple modalities – text and images, learning correlations between them within the conversation. Moreover, the authors created a corpus annotated with intent, symptom, and summary called MM-CliConSumm. In essence, it is a dataset with a doctor impression summary, patient concern summary, and general summary suitable for disease diagnosis. Considering utterance and commonsense knowledge as two different types of data, Xiachong et al. [16]

proposed a heterogeneous dialogue graph network model. It emphasized that explicit knowledge still plays a significant role in summarizing chat conversations.

To handle the lack of paragraph separation in dialogue topic classification, Jiang et al. [30] proposed a self-supervised segmentation model, converting dialogue sentences into semantic vectors with preserved temporal information between dialogues. In other words, the authors have suggested a technique for categorizing topics into groups of paragraphs by understanding various methods of arranging conversational sentences. This method demonstrates the importance of organizing information in a sequential order. Regarding software development projects, Gilmer et al. [26] highlighted a set of guidelines advantageous for collaborative issue discussions summarization tool design in the form of a browser plugin. Their approach aided the process by collaborative authoring and suggesting summaries for user review, editing, and approval using automated summarization techniques. Alternatively, Silva et al. [3] began with manual reflexive thematic analysis of Gitter chat rooms, which involved analyzing automatically generated summaries of chat rooms by hand. The authors applied thematic mapping to compare IM content with the established knowledge framework – SWEBOK, to identify similarities and differences. The outcome established the map of themes suggested for developing data-driven software tools for knowledge identification in IM or collaborative environments for effective developer communication.

The significant outputs of the surveyed articles, distinct or more general summarization challenges addressed by authors, along with manual evaluation and data preprocessing, are summarized in Table 2.

The consistently observed pattern is that the most challenging output is designing a viable application to deliver the summarization to its intended user, apart from different summarization models, which usually cover particular cases. Only four groups of researchers attempted to create a prototype or partially the user interface concept. The latest information indicates that the Collabot tool proposed by Tepper et al. [31] did not persist despite the convenient representation of a web-based application providing personalized visual summaries of group chats. The software has been developed as a chat assistant service that learns users' interests and social ties within a chat group. It provides a digestion of missed content, including main topics discussed, business actors participating, and links and useful resources. The application utilizes the chat channel social graph and their normalized weights in the form of a list of users' connections. Gilmer et al. [26] developed advantageous guidelines for collaborative issue discussion summarization tool design, implicitly targeting software developers working on open-source software projects using issue tracker systems and particularly emphasizing the need for summarization moderation. The authors of the tool emphasized the significant discrepancy between comprehending the practice of knowledge storage and proposing feasible design alternatives to facilitate the summarization process and the optimal utilization of the resulting summaries. Notwithstanding these challenges, the research team succeeded in developing the SUMMIT tool through the conduct of a formative user study. To reduce the necessity for manual input, SUMMIT tool employs automated techniques for identifying information types and summarizing texts, thus supplementing the current GitHub Issues user interface. This functionality enables users to collectively construct, edit, accept, and moderate summaries of different types of information discussed, as well as a set of comments representing continuous conversations within the thread. However, ADSum has also not received attention from the practitioners.

Table 2. Summarization techniques and significant outputs of surveyed articles

Surveyed articles	Distinct and General summarization challenges addressed	Designed Applications, Models, Datasets, and Guidance for summarization	Manual evaluation	Manual Preprocessing
[3]	CT, TM	Map of themes from Gitter	Thematic analysis of chat rooms on Gitter	
[4]	ML, MS, LE	Domain Adapted Model		
[5]	CT, LC	Prototype	Annotation and tagging	✓
[15]	VT	MM-CliConSumm, MM-CliConSummation	Adequacy, fluency domain relevance, consistency, and informativeness	
[16]	UK	D-HGN model	Abstractedness, informativeness, correctness	
[24]	DR, PS, CT	Slack conversation dataset, disentanglement algorithm		
[25]	GD, GT	ADSum tool	Formative study (survey), grammatical fluency, relevance, and accuracy	✓
[26]	LD	SUMMIT tool, Guidelines for design		
[27]	RT, CL	Chatsum – pilot tool study Feature-based model	Annotation for the gold standard	✓
[28]	LS, MS, OOV	Utterances Relation Aware Model		
[29]	FC, MS, DR, CL, LS	FA-DS Model	Factualness, succinctness, and informativeness	
[30]	CT, PS	Model for topic division		
[31]	LP	Collabot – Web-based application		
[32]	WR, BC, CT, MS, CL, RT	TGDGA) model	Relevance, readability	
[33]	LS, FD	Semantic understanding enhanced model with self-supervised methods		
[34]		Chat2Doc chat application		

BC – Backchanneling
CL – Use of abbreviations, acronyms, poor grammar, emoticons, misspelled words
CT – Changing topic
DR – Different roles and perspectives FC – Factual consistency, despite high matching evaluation metrics
FD – Fine-tuning is unstable on small datasets and causes performance degradation
GD – Project maintainers tend to annotate and move **live chat dialogs** to the GitHub Discussions threads manually
GT – No **ground truth summarization exists for Gitter live** chats
LC – Lack of annotated chat corpora for social media applications
LD – Lack of key design and features requirements for summarization application
LE – Lacks semantic evaluation to cover fluency, comparison to human performance
LP – Lack of personalized summarizations
LS – Lack of semantic and structural understanding
ML – Large and complex pre-trained models, the necessity of compressing them
MS – multiple-speakers dialogue
OOV – Out of Vocabulary
PS – Absence of explicit paragraph separation
RT – Real-time nature, synchronous inter-change of utterances in the dialogue
TM – Themes not identified, relation of themes not mapped to SWEBOK
UK – Utterance and commonsense knowledge are not considered two different types of data
VT – Critical significance of visuals in telemedicine
WR – Word repetition – names of interlocutors and important actions

It has been observed that the two industries with the highest demand for quality summarization are medical advising and software development, but the least struggle to find a uniform solution.

This may be due to the complexity of discussion threads, multiple topics, participants with different knowledge levels, interruptions, backchanneling, changing topics, informal language, unusual features, colloquial expressions, and more.

It is important to note that none of the articles, selected or even forming part of the initially identified list of articles, addressed the methods of storing extracted knowledge effectively within organizations or project teams in a way that supports the access of reusable and easy-to-access organizational knowledge, despite storing being essential after the extracting phase. This gap may be addressed through the development of solutions based on structured data representations, including ontologies, knowledge graphs, taxonomies, and semantic models. One of the potential avenues for extracting and transforming unstructured information into machine-interpretable data, with a particular emphasis on its pivotal role in developing sophisticated digital knowledge-based platforms, is the use of ontologies and vocabularies. It is one of the methodologies designed for the manufacturing procedure use case to facilitate extracting and representing procedural knowledge from industrial documents. The most essential aspect of this methodology [36] is the representation of document summaries based on standard vocabularies. A set of ontologies and vocabularies has been employed for the implementation of such representations, comprising annotation, manufacturing, and procedure modules [36]. As a result, the structured representation of procedures, in this manner, allows for the extraction and linking of summaries with other structured representations, thereby enabling the effective sharing and querying of procedural knowledge for subsequent reuse.

A different study conducted research on the analysis of social media discourse and developed a methodology incorporating a proprietary integrated ontology that amalgamates social media metadata with various types of linguistic knowledge, including entities and PropBank role sets [37]. This was then populated with a knowledge graph structure, along with tweets extracted on the topic in question [37], thereby providing a more detailed semantic layer, which is beneficial for comprehending social media discourse.

4.5 Dialogue Summarization Metrics and Evaluation Challenges

The most prominent motive found and proved in 8 articles is the necessity of human evaluation of the summarization output, as they are often produced with high matching metrics (ROUGE) but tend to be inaccurate from the content perspective. The ROUGE metric represents a mathematical approach to evaluation and does not provide a proper semantic comparison. It only provides limited information and does not give a complete picture, such as fluency or comparison to human performance. Hence, we emphasize the inevitability of human evaluation of the proposed model performance evaluation. The authors of [27] attracted medical domain experts for manual annotation of conversations as “to be included in the summary” and “to be excluded” and used these annotations as the gold standard data for training and testing the model, while Tiwari et al. [15] used such metrics as “adequacy,” “fluency domain relevance,” “consistency,” and “informativeness” for manual evaluation. Fan et al. [25] highlighted the lack of labeled data for live chats on Gitter compared to GitHub Discussion data and performed a manual quality evaluation of summarizations considering “grammatical fluency,” “relevance,” and “accuracy.” Considering alternative evaluation methods to BERTScore and FEQA, Li et al. [29] performed human evaluation engaging NLP researchers and evaluated 100 samples to score the “factualness,” “succinctness,” and “informativeness”. Silva et al. [3] were forced to identify by hand software engineering themes based on the description of 87 public developer Gitter chat rooms, and Sinha et al. [5] performed manual discussion threads parsing, annotation, and tagging. Xiachong et al. [16] and Zhao et al. [32] applied manual summarization evaluation using similar metrics: “abstractedness,” “informativeness,” “correctness,” “relevance,” and “readability” respectively.

Referring to research on integrating additional information such as dialogue topics, conversation stages, and key information sequences extracted from the same dialogue, some methods begin with a factual perspective and enhance the credibility of summarization by incorporating preextracted knowledge graphs or external knowledge. However, these methods did not effectively limit or control the effectiveness of knowledge after its addition, so the importance of factual information is only indirectly reflected in the ROUGE metrics [29].

5 Conclusions and Future Work

The main objective of the current survey was to provide insights into the challenges and limitations of techniques used for extracting useful information and summarizing chat conversations. Therefore, useful information extraction and knowledge discovery from a variety of text-based communication mediums have been examined from the perspective of organizational knowledge as a source.

With conclusive evidence, an equally significant aspect of knowledge perspective is its representation in such knowledge-intensive domains as software development and telemedicine. They were distinctly represented as ontologies or commonsense knowledge, suggesting that domain-specific concepts and terms are crucial for quality information extraction from chat conversations and subsequent reuse for productive workflows.

It can be stressed that KA, sharing, creation, and reusable and commonsense knowledge are mutually interrelated. These are concentrated forms of industry or project-specific knowledge, and their cyclical nature is characterized by improved extracted information. Useful information appeared to transform into knowledge by applying knowledge resources and KA, becoming reusable for subsequent knowledge infusion in summarization tasks. The process of transforming useful information into organizational or project knowledge through knowledge discovery and knowledge sharing is iterative. Text-based communication should be perceived as a source of useful information within a project or organization and can be transferred into reusable knowledge represented by KA using knowledge discovery techniques, namely summarization.

To conclude, significant efforts within the scientific community during the last decade have been made to comprehend the role of IM in KM and its potential to enhance organizational productivity and facilitate the learning process, especially in the knowledge-intensive areas. It has been observed that the two industries with the highest demand for quality summarization are medical advising and software development, but the least struggle to find a uniform solution. This is due to the complexity of discussion threads, multiple topics, and participants with different levels of knowledge, as well as interruptions, backchanneling, changing topics, unusual features, and colloquial expressions.

It has been revealed that human evaluation of summarization output is highly demanded due to the limitations of ROUGE metrics, which are primarily used for the evaluation of automatic summarizations but, at the same time, are time-consuming and cost-ineffective. The optimized solution for human-like appraisal reflecting content quality and semantic comparison by such criteria as adequacy, fluency, domain relevance, consistency, informativeness, grammatical fluency, accuracy, factualness, succinctness, correctness, and readability is highly demanded and suggested for future work.

Moreover, there is a need for a uniform user interface and solutions design of a viable application for software development since the industry has the highest demand for quality summarizations to enhance teams' productivity and facilitate the learning process. In addition, there is a lack of larger annotated data sets for multiparty asynchronous chats tailored and adapted to software development domains. Overcoming of this problem is one more topic for further research in chat-based dialog summarization.

References

[1] H. Ajjan, R. Hartshorne, Y. Cao, and M. Rodriguez, "Continuance use intention of enterprise instant messaging: a knowledge management perspective." *Behavior & Information Technology*, vol. 33, no. 7, pp.678–692, 2014. Available: https://doi.org/10.1080/0144929X.2014.886722

[2] G. M. P. Wanderley, C. A. Tacla, J.-P. Barthès, and E. C. Paraíso, "Knowledge discovery in task-oriented dialogue," *Expert Systems with Applications*, vol. 42, no. 20, pp. 6807–6818, 2015. Available: https://doi.org/10.1016/j.eswa.2015.05.005

[3] C. C. Silva, M. Galster, and F. Gilson, "A qualitative analysis of themes in instant messaging communication of software developers," *Journal of Systems and Software*, vol. 192, p. 111397, 2022. Available: https://doi.org/10.1016/j.jss.2022.111397

[4] R. Sroch, "Domain Adapted Abstractive Summarization of Dialogue using Transfer Learning," in *Proceedings of 2021 4th International Conference on Algorithms, Computing and Artificial Intelligence, ACAI'21*, pp. 1–6, 2021. Available: https://doi.org/10.1145/3508546.3508640

[5] A. Sinha, T. K. M. Manohar, S. Subramanian, and B. Das, "Text segregation on asynchronous group chat," *Procedia Computer Science*, vol. 171, pp. 1371–1380, 2020. Available: https://doi.org/10.1016/j.procs.2020.04.147

[6] Y. Levy and T. Ellis, "A systems approach to conduct an effective literature review in support of information systems research," *Informing Science the International Journal of an Emerging Transdiscipline*, vol. 9, pp. 181–212, 2006, Available: https://doi.org/10.28945/479

[7] H. Snyder, "Literature review as a research methodology: An overview and guidelines," *Journal of Business Research*, vol. 104, pp. 333–339, 2019. Available: https://doi.org/10.1016/j.jbusres.2019.07.039

[8] Microsoft Corporation, "Will AI Fix Work?" Work Trend Index Annual Report, 2023. Available: https://www.microsoft.com/en-us/worklab/work-trend-index/will-ai-fix-work. Accessed on December 3, 2023.

[9] Microsoft Corporation, "Microsoft Copilot for Microsoft 365 Overview," 2023. Available: https://learn.microsoft.com/en-us/microsoft-365-copilot/microsoft-365-copilot-overview. Accessed on December 3, 2023.

[10] Google LLC, "Conversation Summaries in Google Chat," 2022. Available: https://blog.research.google/2022/11/conversation-summaries-in-google-chat.html. Accessed on December 17, 2023.

[11] Zoom Video Communications, Inc., "Getting started with Zoom AI Companion features," 2023. Available: https://support.zoom.com/hc/en/article?id=zm_kb&sysparm_article=KB0057623#collapseTeamChat. Accessed on December 2, 2023.

[12] J. Sifakis, *Understanding and Changing the World: From Information to Knowledge and Intelligence*. 1st ed., Springer, 2022. Available: https://doi.org/10.1007/978-981-19-1932-9

[13] C. Schaefer and A. Makatsaria, "Framework of data analytics and integrating knowledge management," *International Journal of Intelligent Networks*, vol. 2, pp. 156–165, 2021. Available: https://doi.org/10.1016/j.ijin.2021.09.004

[14] G. Salazar-Torres, E. Colombo, F. S. C. Da Silva, C. A. Noriega, and S. Bandini, "Design issues for knowledge artifacts," *Knowledge-Based Systems*, vol. 21, no. 8, pp. 856–867, 2008. Available: https://doi.org/10.1016/j.knosys.2008.03.058

[15] A. Tiwari, A. Saha, S. Saha, P. Bhattacharyya, and M. Dhar, "Experience and Evidence are the eyes of an excellent summarizer! Towards Knowledge Infused Multi-modal Clinical Conversation Summarization," in *Proceedings of the 32nd ACM International Conference on Information and knowledge management, CIKM'23*, pp. 2452–2461, 2023. Available: https://doi.org/10.1145/3583780.3614870

[16] F. Xiachong, F. Xiaocheng, and B. Bing, "Incorporating commonsense knowledge into Abstractive Dialogue Summarization via Heterogeneous Graph Networks," in *Proceedings of the 20th Chinese National Conference on Computational Linguistics*, pp. 964–975, 2021. Available: https://doi.org/10.1007/978-3-030-84186-7

[17] Google Cloud, "What is Retrieval Augmented Generation (RAG)? Google," 2024. Available: https://cloud.google.com/use-cases/retrieval-augmented-generation. Accessed on July 3, 2024.

[18] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W. Yih, T. Rocktäschel, S. Riedel, and D. Kiela, "Retrieval-augmented generation for knowledge-intensive NLP tasks," in *Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS'20)*, article 793, pp. 9459–9474, 2020.

[19] Y. Ahn, S.-G. Lee, J. Shim, and J. Park, "Retrieval-Augmented Response Generation for Knowledge-Grounded Conversation in the Wild," *IEEE Access*, vol. 10, pp. 131374–131385, 2022. Available: https://doi.org/10.1109/ACCESS.2022.3228964

[20] M. Jarke, M. A. Jeusfeld, P. Peters, and K. Pohl, "Coordinating distributed organizational knowledge," *Data and Knowledge Engineering*, vol. 23, no. 3, pp. 247–268, 1997. Available: https://doi.org/10.1016/S0169-023X(97)00014-1

[21] C. W. Holsapple and K. Joshi, "Organizational knowledge resources," *Decision Support Systems*, vol. 31, no. 1, pp. 39–54, 2001. Available: https://doi.org/10.1016/s0167-9236(00)00118-4

[22] A. Carneiro, "The role of intelligent resources in knowledge management," *Journal of Knowledge Management*, vol. 5, no. 4, pp. 358–367, 2001. Available: https://doi.org/10.1108/EUM0000000006533

[23] N. Matta, H. Atifi, and F. Rauscher, "Knowledge extraction from professional emails," *2014 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pp. 1407–1414, 2014. Available: https://doi.org/10.15439/2014F161

[24] P. Chatterjee, K. Damevski, N. A. Kraft, and L. Pollock, "Software-related Slack Chats with Disentangled Conversations," *17th International Conference on Mining Software Repositories (MSR), IEEE/ACM*, pp. 588–592, 2020. Available: https://doi.org/10.1145/3379597.3387493

[25] Fan, G., Chen, S., Wu, H., Gao, C., Xiao, J., Xue, X., and Feng, Z., "Dialog summarization for software collaborative platform via tuning pre-trained models," *Journal of Systems and Software*, vol. 204, p. 111763, 2023. Available: https://doi.org/10.1016/j.jss.2023.111763

[26] S. Gilmer, A. Bhat, S. Shah, K. M. Cherry, J. Cheng, and J. Guo, "SUMMIT: Scaffolding open source software issue discussion through summarization," in *Proceedings of the ACM on Human-computer Interaction*, vol. 7, no. CSCW2, article 297, pp. 1–27, 2023. Available: https://doi.org/10.1145/3610088

[27] H. Zafari and F. Zulkernine, "Chatsum: An Intelligent Medical Chat Summarization Tool," in *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, pp. 1–2, 2021. Available: https://doi.org/10.1109/BHI50953.2021.9508585

[28] H. Li, "URAMDS: Utterances Relation Aware Model for Dialogue Summarization: A Combined Model for Dialogue Summarization," in *2nd International Conference on Bioinformatics and Intelligent Computing (BIC 2022)*, pp. 397–401, 2022. Available: https://doi.org/10.1145/3523286.3524577

[29] W. Li, X. Zhou, X. Bai, and S. Pan, "Improving Factual Consistency of Dialogue Summarization with Fact-Augmentation Mechanism," in *International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, 2022. Available: https://doi.org/10.1109/IJCNN55064.2022.9891885

[30] L. Jiang, Y. Hao, and J. Lin, "Dialogue Text Summarization Method Combined Self-supervised Learning and Neural Architecture Search," in *2nd International Conference on Big Data, Artificial Intelligence and Risk Management (ICBAR)*, pp. 65–69, 2022. Available: https://doi.org/10.1109/ICBAR58199.2022.00020

[31] N. Tepper, H. Anat, B. Maya, R. Inbal, and L. Lior, "Collabot: Personalized Group Chat Summarization," in *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, WSDM'18*, pp. 771–774, 2018. Available: https://doi.org/10.1145/3159652.3160588

[32] L. Zhao, W. Xu, and G. Jun, "Improving Abstractive Dialogue Summarization with Graph Structures and Topic Words," in *Proceedings of the 28th International Conference on Computational Linguistics*, pp. 437–449, 2020. Available: https://doi.org/10.18653/v1/2020.coling-main.39

[33] H., Lee, J. Yun, H. Choi, S. Joe, and Y. L. Gwon, "Enhancing semantic understanding with self-supervised methods for abstractive dialogue summarization," in *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, pp. 796–800, 2021. Available: https://doi.org/10.21437/Interspeech.2021-1270

[34] B. Meixner, M. Lee, and S. Carter. "Chat2Doc: From Chats to How-to Instructions, FAQ, and Reports," in *Proceedings of the 2017 ACM Workshop on Multimedia-based Educational and Knowledge Technologies for Personalized and Social Online Training, MultiEdTech'17*, pp. 27–30, 2017. Available: https://doi.org/10.1145/3132390.3132392

[35] H. R. Nemati, C. D. Barko, O. Maimon, and L. Rokach, Eds., “Organizational Data Mining,” *Data Mining and Knowledge Discovery Handbook*, Springer, pp. 1041–1048, 2009. Available: https://doi.org/10.1007/978-0-387-09823-4_55

[36] A. Rula, G. R. Calegari, A. Azzini, I. Baroni, and I. Celino, “Annotation and Extraction of Industrial Procedural Knowledge from Textual Documents,” in *Proceedings of the 12th Knowledge Capture Conference 2023 (K-CAP’23)*, pp. 1–8, 2023. Available: https://doi.org/10.1145/3587259.3627570

[37] I. Blin, L. Stork, L. Spillner, and C. Santagiustina, “OKG: A Knowledge Graph for Fine-grained Understanding of Social Media Discourse on Inequality,” in *Proceedings of the 12th Knowledge Capture Conference 2023 (K-CAP’23)*, pp. 166–174, 2023. Available: https://doi.org/10.1145/3587259.3627557