ICA-CRMAS: Intelligent Context-Awareness Approach for Citation Recommendation based on Multi-Agent System

HOUSSEM EDDINE DEGHA*, Department of Mathematics and Computer Science, University of Ghardaia, Ghardaia, Algeria
FATIMA ZOHRA LAALLAM, Department of Mathematics and Computer Science, Kasdi Merbah Ouargla University, Ouargla, Algeria

Navigating the ever-expanding sea of scientific literature presents a daunting challenge for researchers seeking relevant and up-to-date information. Traditional citation recommendation systems, while well-intentioned, often fall short due to their limited focus on text-based features and lack of contextual awareness. In this paper we introduce the ICA-CRMAS (Intelligent Context-Aware Approach for Citation Recommendation based on Multi-Agent System), an intelligent system that leverages the power of deep learning, semantic analysis, and multimodal learning to overcome these limitations. ICA-CRMAS goes beyond the surface, delving into the rich tapestry of information within academic papers, including figures, which often hold vital contextual clues. By weaving this contextual data directly into its recommendation models, ICA-CRMAS generates highly personalized and relevant suggestions. This comprehensive approach unlocks enhanced accuracy, diversity, and serendipity, enabling researchers to effectively discover papers aligning with their interests and research objectives. ICA-CRMAS illuminates its reasoning. Instead of opaque suggestions, the system provides clear explanations that justify and illustrate recommended citations. This transparency builds user confidence, allowing researchers to critically engage with and trust the system's recommendations. Evaluation experiments conducted on real-world academic datasets demonstrate that ICA-CRMAS outperforms existing approaches across various metrics. it surpassing its closest competitor by a margin of 7.53 on accuracy, 6.07% on MRR and by 5.87 on Recall. User feedback further reinforces its effectiveness, with an Overall System Usability Scale (SUS) score of 76.73, exceeding benchmark scores for comparable systems.

CCS Concepts: • **Information systems → Recommender systems**; *Information extraction*; Query representation; *Data analytics*; *Content ranking*; • **Applied computing → Computer-assisted instruction**.

Additional Key Words and Phrases: Multi-agent Systems, Ontology, Scientific Papers, Recommendation Systems, Search Engines, User Profiles, Personalized Information Filtering, World Wide Web, Context-awareness

1 Introduction

The exponential growth of the internet and the rise of numerous journal publishers have led to a substantial increase in the volume of published scientific literature. Modern scientific digital libraries (DLs) now curate vast collections of scholarly works, comprising millions of papers and billions of references. Scopus, a prominent scholarly database, boasts over 71 million publications and 1.4 billion references. Web of Science, another comprehensive platform, houses over 78 million records in its core collection and over 171 million records across the Web of Science platform . Various DLs and publishers, including Nature Research, the Association for Computing

Authors' Contact Information: Houssem Eddine Degha, Department of Mathematics and Computer Science, University of Ghardaia, Ghardaia, Ghardaia, Algeria; e-mail: degha.houssem@outlook.com; Fatima Zohra LAALLAM, Department of Mathematics and Computer Science, Kasdi Merbah Ouargla University, Ouargla, Ouargla, Algeria; e-mail: laallamfz@gmail.com.

ACM 2158-6578/2024/8-ART
https://doi.org/10.1145/3680287

Machinery, Springer, IEEE Transactions, Elsevier, DOAJ, and JSOR, also play crucial roles in disseminating high-quality scientific papers.

A plethora of scholarly search engines, including Google Scholar, Microsoft Academic, Science.gov, and Semantic Scholar, are widely utilized for research purposes [1, 28]. Google Scholar, a freely accessible web search engine, hosts over 389 million records, as estimated by Michael [28]. Microsoft Academic, another prominent option, is a free public web search engine for academic publications, covering 241 million publications as of 2020 [1]. The vast and ever-growing repository of scientific literature presents a challenge for researchers, particularly those new to their field, who often struggle to identify relevant papers using effective keywords. A significant obstacle faced by master's and Ph.D. students is the overwhelming volume of scientific papers returned after executing a query. Most search engines yield thousands or even millions of papers, making it time-consuming and challenging for users to manage effectively. While existing scholarly search engines employ intelligent algorithms and a variety of quality metrics to evaluate and rank scientific papers, this issue persists. Quality metrics, such as Impact Factor, H-index, g-index, consistent annual citation-based index, s-index, and topic-based metrics, play a crucial role in assessing and recommending recent high-quality scientific papers [15]. However, the sheer volume of literature makes it difficult for researchers to efficiently identify the most relevant and impactful papers among the vast sea of available research.

In the realm of scientific literature, researchers and students grapple with the overwhelming volume of information, necessitating solutions to enhance search efficiency. The transformation of manual web searches into semi-automatic or automatic tasks emerges as a key solution, with recommendation systems (RS) standing out as a promising approach. Empowered by big data and data analytics, RS achieves user satisfaction through personalization by selecting and proposing relevant items based on user profiles, interests, preferences, and needs.[27]. While recommendation systems offer significant benefits, they encounter challenges. The Cold-start problem arises with new users or items, impeding the system's ability to provide relevant recommendations. Data Sparsity, particularly in extensive catalogs, hampers the identification of relevant papers. Additionally, low accuracy becomes a concern when the system struggles to predict users' preferences accurately. Recent advancements in machine learning and deep learning have paved the way for addressing the shortcomings of traditional citation recommendation techniques and enhancing their accuracy [19, 38]. In 2020, Ali et al. conducted a comprehensive analysis of these models, categorizing them based on factors such as data utilization, representation methods, recommendation services, and personalization strategies [4]. Their findings revealed that incorporating information such as paper contents, tags, user profiles, and citation networks can effectively represent researchers' interests and enhance the relevance of recommendations.

Despite the progress made in citation recommendation methods, a notable research gap exists in the existing literature. This gap stems from several significant limitations in current approaches. Firstly, many of these methods heavily rely on text-based features from academic papers while overlooking valuable information present in figures, tables, and other visual elements. This narrow focus limits the comprehensive understanding of scholarly documents, thereby hindering the ability to offer truly holistic recommendations. Secondly, the lack of explainability and transparency in existing citation recommendation techniques is another critical issue. This opacity can lead to a lack of trust among users, potentially reducing their willingness to adopt the recommendations. Without clear insight into how recommendations are generated, users may perceive them as arbitrary or irrelevant. Lastly, current approaches primarily focus on low semantic levels, such as the paper-citation relationship, when identifying similarities between papers. While this approach provides a basic level of understanding, it falls short in capturing the nuanced contextual information and complexities inherent in scholarly works. A system that can effectively comprehend and leverage intricate contextual cues can operate at higher semantic levels, thereby offering more accurate and meaningful citation recommendations. Addressing these gaps requires a shift towards contextual approaches in recommendation systems. These approaches aim to enhance accuracy by incorporating contextual information from various sources, such as user profiles, document characteristics, and environmental factors. Context-Aware Recommendation Systems (CARS) utilize this information to deliver personalized and relevant suggestions. However, integrating contextual information into recommendation systems presents new challenges. It is essential to accurately represent and utilize this information in a precise and traceable manner for effective implementation of CARS. Additionally, establishing context-awareness and leveraging it to enhance recommendation accuracy necessitates careful consideration and experimentation in future research efforts.

To address the challenges and limitations of existing methods, we propose ICA-CRMAS (Intelligent Context-Awareness approach for citation recommendation based on Multi-Agent System). ICA-CRMAS employs an intelligent reasoning mechanism based on a context-awareness approach to provide personalized and accurate citation recommendations. It distinguishes itself by utilizing multimodal learning to harness information from various modalities, including text, figures, and tables, for a nuanced understanding of the citation context. It also leverages its ability to handle complex contextual information, accessing higher semantic levels for more refined recommendations. Additionally, ICA-CRMAS enhances explainability and transparency by providing clear justifications and visualizations for its recommendations, fostering user trust and adoption. Overall, ICA-CRMAS aims to revolutionize citation recommendation by offering accurate, personalized, and trustworthy suggestions tailored to researchers' specific needs and interests.

The paper progresses through several sections to comprehensively explore the research topic. First, the literature review presents an overview of pertinent studies, laying the foundation for subsequent sections. Next, the "ICA-CRMAS Architecture" section details the proposed system's components and functionality. Following that, the "Development" section outlines the implementation process, emphasizing the techniques and tools employed. Then, the "Experiments & Evaluation" section applies the system to a specific dataset and provides a rigorous evaluation of its performance. This demonstrates the practical value of ICA-CRMAS and its potential impact on researchers. Furthermore, the "Illustrative Examples" section furnishes concrete examples, while the "SUS Evaluation" section assesses user satisfaction and usability. Finally, the conclusions summarize the findings, discuss the research contributions, and suggest future directions for further exploration.

2 Related works

Significant advancements have characterized the field of citation recommendation in recent years, fueled by innovative contributions from various research endeavors [23]. Recommendation systems (RS) constitute a multi-faceted domain encompassing diverse technologies and methodologies designed to deliver relevant information to users. These systems excel at processing vast amounts of data, particularly textual information. Key categories of recommendation technologies include: Content-Based Filtering (CBF), Collaborative Filtering (CF), Graph-Based Filtering (GBF), Context-Based Techniques and Hybrid Recommender Systems.

Content-based citation recommendation techniques analyze textual content to generate suggestions using similarity measures like cosine similarity to rank candidate citations. Several studies have applied these methods. Chandra et al. [5] in 2018 developed a citation recommendation method that embeds paper content in vector space and employs machine learning to refine suggestions, releasing an online portal for citation recommendations. In 2020, Kieu et al. [34] used Sentence-BERT for recommending scientific literature, achieving better results than other approaches on the ACL Anthology Network corpus using metrics like MRR and F1@k score, with performance improvements of 3.67% to 6.56%. Guo et al. [27] in 2021 introduced CSCR, which combines Doc2vec embeddings with network enhancement to improve recommendation accuracy and efficiency in vast bibliographic networks, showing significant performance boosts over benchmark methods. Content-based techniques are simple and efficient, not requiring external knowledge. However, they face challenges like vocabulary mismatch and may miss nuanced intent and influential factors in citation decisions. The limitations include the need for detailed item descriptions, complementary recommendations rather than implied content, focus on individual user preferences without serendipity, and tendency to suggest items similar to previous encounters.

Collaborative Filtering (CF) systems filter information by leveraging users' historical interests, offering several advantages. They exploit collective preferences of large user groups, provide recommendations for both complementary and substitute items, introduce users to compelling items they may not actively seek, and facilitate knowledge exchange among users with similar interests. Several researchers have investigated CF techniques. In 2015, Lee et al. developed a CF-based individual recommendation system for researchers, employing the Simple K-Nearest Neighbors algorithm and a text vectorization technique for paper representation, showing positive user satisfaction but limited to three participants [64]. Similarly, X. Kong et al. proposed a citation recommendation method relying solely on detectable citation relations, outperforming traditional CF methods but lacking incorporation of authors' social relations [37]. In 2020, Nazmus et al. enhanced scientific paper recommendation systems using a 2-level paper citation relations approach, demonstrating significant precision improvements compared to baseline approaches without relying on prior user profiles [45]. Despite these advantages, CF has limitations, such as dependence on substantial user data for optimal functionality, challenges with the cold start problem for new users, difficulties accommodating users with distinctive preferences, and issues in clustering and classifying users with evolving preferences.

Graph-Based Filtering (GBF) techniques exploit the network structure of scientific literature to generate citation recommendations by viewing the literature as a graph where nodes represent publications and edges represent citations. They use methods like link analysis, path analysis, or network embedding to measure the relevance of candidate citations. Scholars have investigated these techniques extensively. In 2019, XIAO et al. developed a personalized recommendation technique based on heterogeneous graph representation, initializing node vectors with word embedding and selecting recommended papers using meta-path-based proximity measures, showing effectiveness on the DBLP network [39]. In 2020, Chanwoo et al. introduced FullTextPeerRead, a dataset for context-aware paper citation recommendation, and developed a model using BERT and graph convolution networks, achieving significant improvements in mean average precision and recall@k over previous methods. In 2023, Pham et al. discussed the limitations of current academic networks and proposed TopCite, an approach combining content-based and citation relationship-based analyses within a biographical network, demonstrating effectiveness on the AMiner dataset [43]. GBF techniques effectively harness the relationships among publications, but they may not encapsulate the content and context of citations, struggle with the dynamic nature of scientific literature, and face scalability and efficiency challenges with large graphs.

Hybrid techniques strategically integrate multiple methodologies to generate citation recommendations, leveraging the strengths and mitigating the weaknesses of individual techniques. Scholars have extensively investigated these approaches. In 2015, Meilian et al. proposed the AHITS algorithm using a tripartite graph to evaluate the quality and authority of academic resources, effectively recommending high-quality papers but leaving the cold start problem unresolved [42]. In 2017, Ebesu et al. introduced CiteRNN, a tool combining content-based and network-based models to provide citation recommendations by learning joint representations of citation and candidate papers as vectors [21]. In 2020, Hui et al. proposed a hybrid approach using the AMHG graph to address the cold start problem by incorporating paper metadata and author influence, demonstrating enhanced recommendation accuracy on the DBLP-REC dataset, despite relying on offline data and lacking personalized recommendations [30]. Hybrid techniques surpass individual methodologies by combining diverse information, but they also introduce complexity, redundancy, compatibility challenges, and require meticulous parameter tuning.

After conducting an exhaustive review and analysis of various models employed in citation recommendation, we have identified the most pivotal models. A comprehensive comparison was conducted based on four criteria: (1) adopted methodologies, (2) utilized information, (3) data representation models, and (4) addressed problems, as illustrated in Table 1 (For a more comprehensive comparison, refer to Table 13 in Appendix A.). This comparative analysis reveals several shortcomings in existing citation recommendation methods. Current systems often overlook non-textual features such as visual elements, which provide vital contextual information and insights into research topics. Additionally, previous methods inadequately integrate contextual information, typically incorporating it only in pre-filtering or post-filtering stages rather than directly within recommendation models. This limited integration hampers a comprehensive understanding of the citation context, potentially leading to less accurate and relevant recommendations. A final limitation of current citation recommendation approaches is their diminished interpretability and explainability, particularly those utilizing complex machine learning and deep learning models. Researchers often require insights into the rationale behind recommended citations to ascertain their relevance and credibility. However, many systems fail to provide any explanation or justification for their recommendations or only furnish superficial or vague information. This lack of explainability may erode user trust in the system's recommendations.

To address these limitations, citation recommendation approaches need to adopt more comprehensive and multimodal methods that leverage various types of information from academic papers and their contexts. Moreover, enhancing context awareness and ensuring interpretability and explainability will elevate the quality and utility of citation recommendations for researchers. In this paper, we introduce ICA-CRMAS, an Intelligent Context-Awareness approach for citation recommendation based on a Multi-Agent System. Our proposed method seeks to address the aforementioned issues by incorporating an intelligent reasoning process rooted in context awareness. We introduce a groundbreaking approach that sets itself apart from existing methods in several key aspects. Firstly, our approach is grounded in multimodal learning, a novel and promising technique for citation recommendation. Multimodal learning can leverage various modalities, including text and figures, present in the citation recommendation scenario. This methodology enhances the representation and comprehension of the citation context and candidate citations, thereby providing more accurate and diverse recommendations Moreover, ICA-CRMAS incorporates contextual information not only in pre-filtering and post-filtering but also directly into the recommendation models. This comprehensive integration ensures a more accurate understanding of the citation context and facilitates the generation of appropriate recommendations.Thirdly, our approach is rooted in an explainable and transparent perspective, aspects frequently lacking or unclear in existing techniques. Through the utilization of information from multiple modalities, our approach furnishes explanations and visualizations for recommendations, justifying and illustrating the suggested citations. This enhancement promotes user understanding and confidence in recommendations, offering more informative and persuasive suggestions.

	Model	Filtering method								Information used								Data model			Problem faced		
		Collaborative	Content-based	Graph-based	Machine Learning	Deep learning	Multimodal learning	Hybrid	Context-aware	Paper contents	Tags/Keywords	Author profile	Venue information	Citation network	Social network	Rating	Images	Matrix-based	Graph-based	Ontology-based	Sparsity	Cold-start	Accuracy
21	HRLHG ([32])	-	-	✓	-	✓	-	-	-	✓	✓	✓	-	✓	-	-	-	-	✓	-	-	-	-
28	CITEWERTs([25])	-	-	-	✓	-	-	-	✓	✓	-	✓	-	-	-	-	-	✓	-	-	-	-	✓
31	GAN-HBNR ([9])	-	-	✓	-	✓	-	-	-	✓	-	✓	-	✓	✓	-	-	-	✓	-	✓	-	-
32	PCCR ([52])	-	✓	-	-	✓	-	-	-	✓	-	✓	-	✓	-	-	-	-	✓	-	-	-	-
37	VCGAN ([54])	-	-	✓	-	✓	-	-	-	✓	-	✓	-	-	✓	-	-	-	-	-	-	-	-
38	TMR-PCR ([8])	-	-	✓	-	✓-	-	-	-	-	✓	-	✓	-	-	-	-	✓	-	-	-	-	✓
42	CPA-CE ([33])	-	-	-	-	-	-	-	-	✓	-	-	-	✓	-	-	-	-	-	-	-	-	-
30	NNRank ([5])	-	✓	-	-	✓	-	-	-	✓	-	✓	✓	-	✓	-	-	✓	-	-	-	-	✓
8	Ma(2019) [39]	✓	-	-	-	✓	-	-	-	✓	✓	✓	✓	-	-	-	-	-	✓	-	-	-	✓
12	CIRec ([13])	✓	-	-	-	-	-	-	-	✓	✓	-	-	✓	-	-	-	-	✓	-	-	-	-
16	HGRec ([40])	✓	-	✓	-	✓	-	-	-	✓	✓	✓	✓	✓	✓	-	-	-	✓	-	-	✓	-
17	POLAR ([20])	-	-	-	-	✓	-	-	-	✓	-	✓	-	-	-	-	-	✓	-	-	-	-	-
27	ASL ([18])	-	✓	-	-	✓	-	-	✓	✓	-	✓	-	✓	-	-	-	-	✓	-	-	-	-
29	WHIN-CSL ([12])	-	-	✓	-	✓	-	-	-	✓	-	✓	✓	✓	✓	-	-	-	✓	-	-	-	-
35	VOPRec ([35])	-	-	-	✓	-	-	-	-	-	-	✓	-	✓	-	-	-	✓	✓	-	-	-	✓
40	BNR ([11])	-	-	✓	-	✓	-	-	-	✓	-	✓	-	✓	-	-	-	✓	✓	-	-	-	-
41	HRM ([36])	-	✓	-	-	-	-	-	-	✓	✓	✓	✓	-	✓	-	-	✓	-	-	✓	✓	-
43	AED ([50])	-	✓	-	-	✓	-	-	-	✓	-	-	-	✓	✓	-	-	-	-	-	-	-	-
44	NREP ([51])	-	-	✓	-	✓	-	-	-	-	-	✓	-	-	✓	-	-	-	✓	-	-	-	-
45	HIPRec ([41])	-	-	✓	-	-	-	-	-	-	✓	✓	✓	✓	-	-	-	-	✓	-	-	✓	✓
46	CIR ([14])	-	-	✓	-	-	-	-	-	✓	✓	-	-	✓	-	-	-	-	✓	-	-	-	✓
9	Sakib(2020) [45]	✓	-	-	-	-	-	-	-	✓	-	✓	-	✓	-	-	-	-	✓	-	-	-	✓
11	Jeong(2020) [31]	✓	-	-	-	✓	-	-	-	✓	-	-	✓	-	-	-	-	-	✓	-	-	-	✓
18	RI-PR ([7])	-	-	-	-	✓	-	-	-	✓	✓	✓	-	-	-	-	-	-	-	-	-	-	-
19	DocCit2Vec ([53])	-	-	-	-	✓	-	-	-	✓	-	-	-	✓	-	-	-	✓	-	-	-	-	-
48	HybridCite ([24])	-	-	✓	-	✓	-	✓	✓	✓	-	-	-	-	-	-	-	✓	-	-	-	-	-
49	ADRCR ([29])	✓	-	-	-	-	-	-	-	✓	-	-	-	✓	-	-	-	-	-	-	-	-	-
50	GMCT ([48])	✓	-	✓	-	✓	-	-	-	✓	✓	-	-	✓	-	-	-	-	✓	-	-	-	-
51	DER ([16])	-	-	-	-	✓	-	-	-	✓	-	-	-	-	-	-	-	✓	-	-	-	-	-
52	ConvCN ([44])	-	-	-	-	✓	-	-	-	✓	-	-	-	✓	-	-	-	✓	✓	-	-	-	✓
53	GLNNR ([10])	-	-	✓	✓	-	-	-	-	✓	-	-	-	✓	-	-	-	✓	-	-	-	-	-
54	SciNCL ([68])	-	-	✓	✓	-	-	-	-	-	-	-	-	✓	-	✓	-	-	-	-	-	-	-
55	DisenCite ([47])	-	-	✓	-	-	-	-	✓	✓	-	-	-	✓	-	-	-	-	✓	-	-	-	-

Table 1. Comparison between citation recommendation models.

3 ICA-CRMAS Architecture

This section introduces ICA-CRMAS (Intelligent Context-Awareness System for Citation Recommendation based on a Multi-Agent System), a cutting-edge methodology designed to revolutionize citation recommendation. By leveraging multimodal learning, semantic analysis, and user-specific tailoring, ICA-CRMAS offers a holistic framework that promises to significantly enhance the quality, relevance, and personalization of citation recommendations.

The ICA-CRMAS architecture (Figure 1) comprises seven core modules:

(1) **Graphic User Interface**: Facilitates user interaction.
(2) **Query Modeling Module**: Identifies missing information in the data and prepares queries for online retrieval.
(3) **Searching Module**: Utilizes intelligent agents to retrieve missing information from online sources.
(4) **Ontology Module (Onto-CCR)**: Provides a structured framework for organizing data related to citation recommendations.
(5) **Intelligent Context Management (ICM)**: transforms the low-level data into a high-level contextual data representation using ontology to enhances system context awareness.
(6) **Filtering Module**: Employs deep learning algorithms by intelligent agents to recommend highly relevant citations to users.
(7) **Database**: Stores and manages all system data.

ICA-CRMAS integrates contextual information throughout its entire process, from pre-filtering and post-filtering to directly within its recommendation models. This holistic approach ensures the generation of highly accurate and relevant recommendations tailored to each user's individual needs. However, incorporating contextual information presents significant challenges. First, the system must model the citation context in a way that fosters context-awareness. Second, it must effectively utilize this context-awareness to deliver more accurate and relevant recommendations. Compounding these challenges are shortcomings in existing datasets. Each dataset lacks essential data features necessary for building robust contextual citation recommendation models. For instance, the DBLP dataset lacks citation context information, while the ACL Anthology doesn't include publication years for its papers (For dataset specifications, please refer to Appendix B). These deficiencies lead to incomplete and weak data models for citation recommendation. To address these issues, the ICA-CRMAS architecture employs two crucial sub-models. The first module, called Query Modeling, detects all relevant missing information in the dataset and prepares queries to search for the missing information on the web. The second module, the Searching Module, consists of intelligent agents responsible for searching for missing information on the web using queries from the Query Modeling Module.

These processes gather a massive amount of data that requires processing and organization to enrich the existing data structure. Onto-CCR, an Ontology Model, provides the structural framework for this organization. Onto-CCR includes machine-interpretable definitions of fundamental concepts and their relationships, structuring all data in the system to aid the reasoning process and enable the creation of a well-organized contextual data model. The pre-filtering stage utilizes the ICM module, a crucial component of ICA-CRMAS. This module transforms raw data into sophisticated contextual information, elevating context awareness and empowering the system to achieve a deeper understanding.Unlike previous methodologies that focus on processing low-level data representations, ICA-CRMAS emphasizes building awareness within the system by constructing robust high-level data representations. This awareness is then leveraged to enhance citation recommendations. Specifically, the ICM identifies the most pertinent context for utilization by the Filtering module.

The Filtering Module (FM) comprises intelligent agents employing deep learning algorithms that collaborate to recommend the most relevant citations to the user. FM leverages a multimodal learning technique that utilizes various modalities, including text, figures, and all available contextual features, to enhance the representation and comprehension of the citation context and candidate citations. This methodology results in more accurate and diverse recommendations. Furthermore, FM incorporates semantic rules alongside its deep learning core, enabling the filtration of the most relevant citations for the user. This innovative approach integrates context-awareness into the learning process, allowing the system to enhance its understanding and achieve higher accuracy. Our approach prioritizes explainability and transparency, aspects often lacking or unclear in existing techniques. ICA-CRMAS provides explanations by presenting several metrics and deploying all relevant semantic rules used to justify and illustrate the suggested citation. This enhancement promotes user understanding and confidence in recommendations, offering more informative and persuasive suggestions.

ICA-CRMAS has been modularized to minimize intrusiveness of user interaction and maximize flexibility, scalability, extensibility, and interoperability. The architecture of ICA-CRMAS is shown in Figure 1. Overall, ICA-CRMAS offers a novel and comprehensive approach to contextual citation recommendation, addressing the shortcomings of existing techniques and providing a more accurate, relevant, and transparent recommendation experience for users. In the following subsections, we will elaborate on the details of each module in our system and elucidate its operational mechanisms. This will effectively showcase the innovative aspects of our approach (Refer to Appendix C for illustrative examples of ICA-CRMAS application).

3.1 Ontology Model

The utilization of ontologies is a powerful tool for representing knowledge domains. Within this section, we delve into the details of our ontological knowledge model, named Onto-CCR (Contextualized Citation Recommendation Ontology). Designed to provide a structured framework for organizing citation recommendation data, Onto-CCR encapsulates machine-interpretable definitions of key domain concepts and their interconnections.

Onto-CCR encompasses various conceptual entities, including papers, authors, contexts, citations, time, topics, and context-awareness, among others. Each entity is instantiated as a class within our ontology, equipped with data type properties that clarify their characteristics and information attributes. Additionally, each class includes an object property that signifies its association with other classes. Adopting a formal context model grounded in ontologies enhances reasoning capabilities by systematically representing domain knowledge (Figure 2 illustrates key Onto-CCR classes and their relationships). The meticulously structured relationships between concepts within Onto-CCR form the foundation for facilitating reasoning processes (Refer to Appendix D for a comprehensive overview of the essential Onto-CCR classes and their interrelationships).

3.2 Query Modelling Module

The Query Modeling Module, a crucial sub-module within ICA-CRMAS, is responsible for identifying and formulating search queries to acquire missing information. This step is essential for enriching the data structure and ensuring accurate and relevant citation recommendations. The module operates by comparing the existing information of individual models within the system against the base definitions specified in the Onto-CCR ontology. By identifying discrepancies and missing elements, the module generates a list of individuals requiring further data acquisition. Each individual is represented as a node within the ontology graph, facilitating efficient reasoning and prioritization. To determine the most relevant incomplete information for retrieval, the module

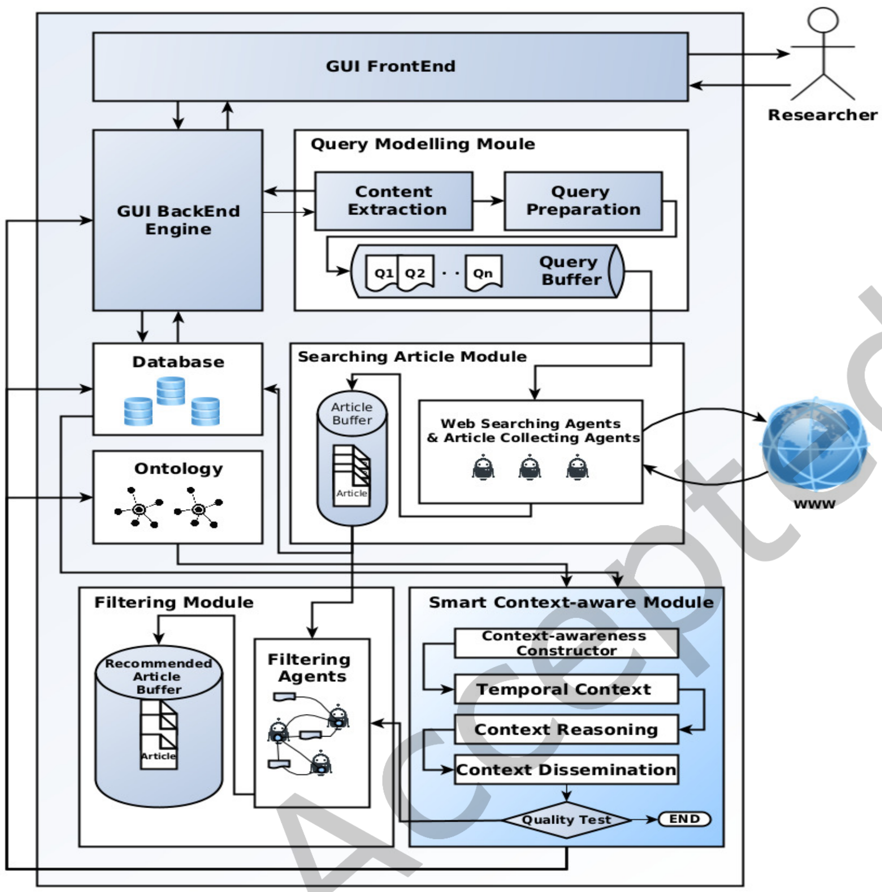


Fig. 1. ICA-CRMAS architecture.

employs the betweenness centrality formula (Equation 1).

$$\begin{aligned}
&\textbf{Node_Power (Ind)} = Normmal(BC(Ind)) * ICP(Ind) \\
&Where: \\
&\textbf{BC(ind)} = \sum_{s \neq Ind \neq t} \frac{\sigma_{st(Ind)}}{\sigma_{st}} \\
&\textbf{Normmal(BC(Ind))} = \frac{(BC(ind) - Min(BC))}{(Max(BC) - Min(BC))} \\
&\textbf{ICP(Ind)} = \frac{NMI}{TNI}
\end{aligned} \tag{1}$$

- **Ind** : `Ontology individual`

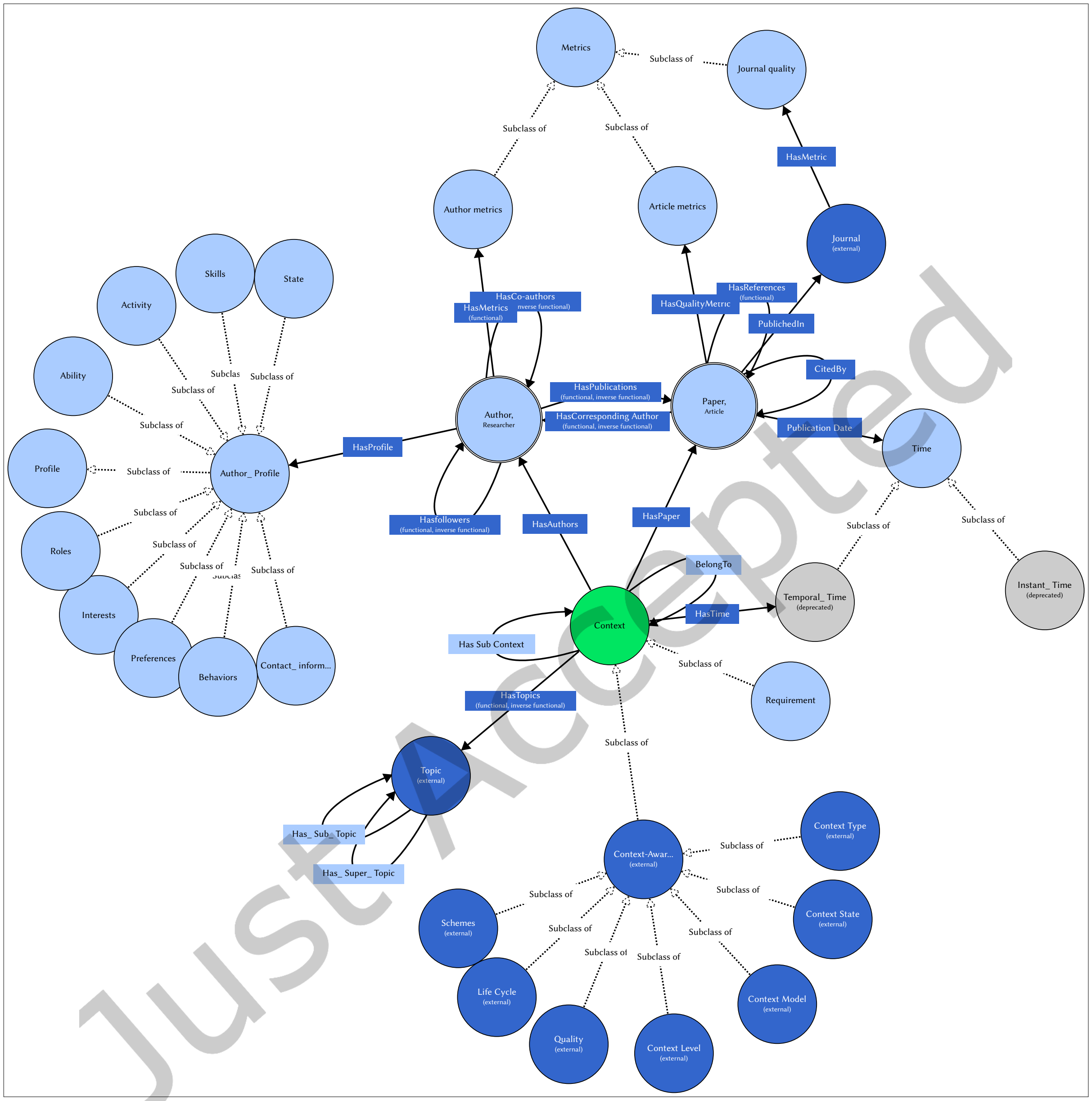


Fig. 2. An illustration of several Onto-CCR ontology concepts and class relationships.

- **Normmal** : It’s the function that normalize the power of nodes and scale it between 0 and 1 without losing precision
- **BC** : Betweenness Centrality
- σ_{st} : It is the total number of shortest paths from node s to node t

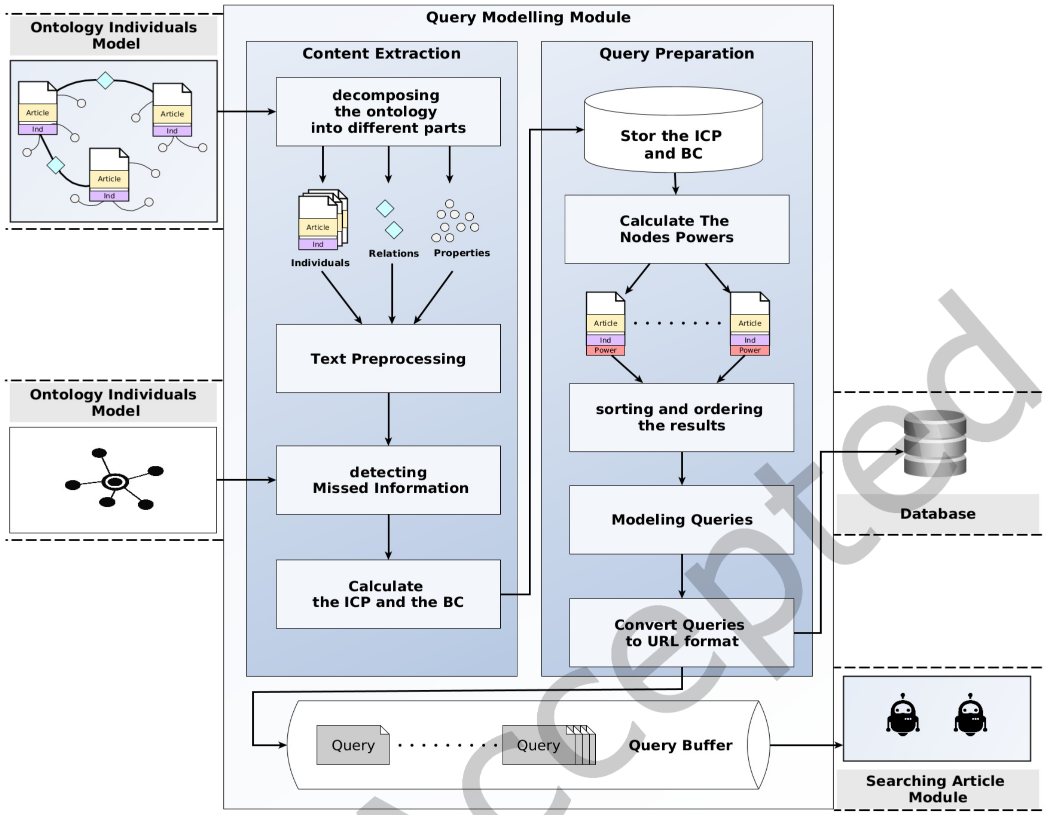


Fig. 3. Represents the sub-architecture of the Query Modelling Module in ICA-CRMAS.

- $\sigma_{st(Ind)}$: It is the number of paths that pass through Ind.
- **ICP** : The probability of completeness of the information
- **NMI** : Number of Missed Information
- **TNI** : Total number of information

This formula calculates the power of each node in the ontology graph based on its centrality and potential impact on the overall recommendation accuracy. Only nodes exceeding a user-defined importance threshold are deemed critical and prioritized for further action. For each prioritized individual, the Query Modeling Module generates search queries targeting the missing information. These queries are then converted into URL format and stored in the query buffer, enabling subsequent retrieval by the Searching Module (For an illustrative example of how the Query Modelling Module handles Detection and Retrieval of Missing Data, please refer to Appendix E). This process allows the system to gather information from diverse academic sources, such as Elsevier, Springer, and Google Scholar, enriching the data structure and facilitating the generation of accurate and context-aware citation recommendations. Figure 3 illustrates the internal architecture of the Query Modeling Module, providing a visual representation of its data flow and information processing steps.

3.3 Searching Articles Module

The Searching Articles Module (SAM) acts as ICA-CRMAS' diligent web explorer, systematically traversing online sources to uncover missing data that completes the Onto-CCR ontology models. This meticulous process unfolds in several key stages:

Web Crawling with Targeted Precision: Guided by targeted queries from the Query Buffer, intelligent agents resembling digital spiders crawl the web. They expertly navigate page labyrinths, extracting relevant URLs that hold the keys to the missing data points. To ensure efficient and agile exploration, these agents leverage the robust JSOUP Java library. This user-friendly tool acts as a map, enabling swift and precise data extraction through fetching URLs, parsing HTML elements, and applying CSS selectors. One of the significant challenges lies in the vast and diverse landscape of web sources, each requiring specific search and collection algorithms. To tackle this complexity, we have developed over 12 specialized agents, each meticulously crafted to navigate and extract data from a specific platform, such as IEEE, ACM, Elsevier, Springer, and many others. This tailored approach ensures efficient and accurate knowledge acquisition across the varied terrain of the web. Once relevant URLs are identified, specialized article-collecting agents take center stage. With meticulous precision, they crawl the identified pages, diligently extracting and scraping information related to the missing data points. This captured knowledge is then transformed into a well-structured format and deposited in the Data Buffer, a temporary repository for these newly acquired treasures. Finally, the extracted information from the Data Buffer is used to update and enrich individual models within the Onto-CCR ontology. This crucial step closes the loop on information acquisition, ensuring that ICA-CRMAS operates on a comprehensive and accurate foundation. Figure 4 provides a deeper look into the intricate workings of the Searching Articles Module, revealing the interconnected interplay between web crawling, data extraction, and knowledge integration.

3.4 Intelligent Context Management (ICM)

The Intelligent Context Management (ICM) module plays a pivotal role in ICA-CRMAS by enriching citation recommendations with contextual understanding. This section delves into the ICM's internal workings, outlining its four key phases that govern the context life cycle:

3.4.1 Context-Awareness Construction. This stage focuses on identifying and modeling relevant contextual information. It begins by detecting high-powered topics nodes, serving as the cornerstone for building individual context models. Each topic model encompasses both static (persistent) and dynamic (evolving) elements, culminating in a comprehensive representation of the topic's contextual landscape.

- **Topic Detection and Modeling:** High-powered topics are identified, and dedicated context models are constructed for each.
- **State Assignment:** Each context is assigned a specific state (Active, Suspended, Resumed, Expired, Terminated) to guide their processing in subsequent phases.
- **Quality-of-Context (QoC) Evaluation:** Employing metrics like author reputation, paper quality, and data validity, a QoC score is assigned to each context, facilitating the selection of the most pertinent information.
- **Context Level Designation:** Contexts are assigned "levels," indicating their degree of processing. Level 0 represents raw information, while higher levels reflect the incorporation of additional reasoning and refinement.
- **Model Integration:** New contexts are either integrated into the existing ontology model or linked as supplemental resources, enriching the overall knowledge base.

Once the context is created, each topic context is linked to papers that share similar keywords, indicating that the paper is related to this specific topic. This linking of paper nodes leads to the connection of all other nodes that

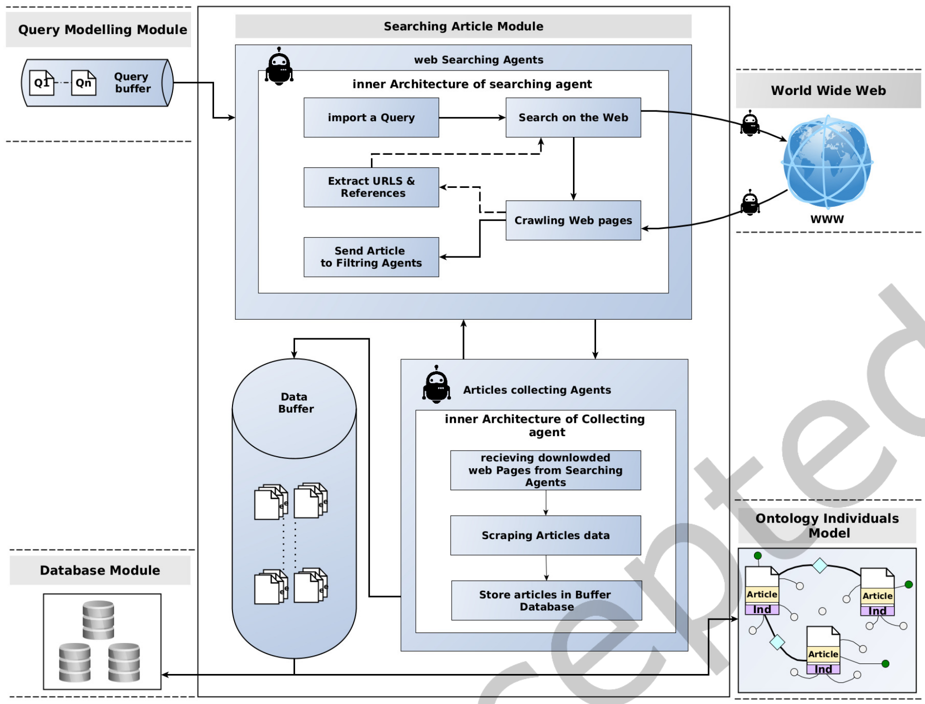


Fig. 4. Illustrates the sub-architecture of the Searching Articles Module in ICA-CRMAS.

have a relation with that paper to the topic context. This process allows the creation of a sub-model representing the contexts.

3.4.2 Temporal Context Processing. This phase optimizes reasoning efficiency by pre-processing the context models generated in the previous stage. Key functions include:

- **Data Cleaning:** Removal of irrelevant or redundant information enhances processing efficacy.
- **Information Merging:** Data obtained from diverse sources is consolidated, ensuring consistency and completeness.
- **Context Prioritization:** Based on QoC scores, significant contexts are prioritized for further processing, streamlining the reasoning process.
- **Temporal Buffering:** Unprocessed contexts are efficiently managed in a dedicated buffer, employing scheduling algorithms (e.g., First-Come-First-Serve) to ensure orderly processing.

3.4.3 Context Reasoning. This stage employs Semantic Web Rule Languages (SWRL and SQWRL) to selectively identify relevant and high-quality candidate citations, offering additional context and insights to enhance citation relevance. Through the utilization of ontology and semantic rules, our objective is to elevate the accuracy and diversity of citation recommendations (For examples illustrating the implementation of rules in Context Reasoning, see Appendix F). The ICM module plays a crucial role in transforming low data level into a higher contextually aware knowledge level, thereby empowering ICA-CRMAS to generate precise and insightful citation recommendations. Its four-phase life cycle ensures efficient context management, leading to a comprehensive understanding of the citation landscape and ultimately enhancing the system's effectiveness.

3.5 Filtring Article Module

This section offers a detailed exploration of the Filtering Article Module (FAM)'s internal architecture. We delve into the specific algorithms and techniques employed in each component, granting insights into the functionality of FAM and its capacity to filter and recommend highly relevant papers to users based on their queries. The FAM architecture operates through three distinct stages: retrieval, ranking, and post-ranking. By meticulously examining each stage, this section aims to provide a comprehensive understanding of FAM's operation and its crucial role in delivering precise paper recommendations. Prior to embarking on the filtering process, the FAM undergoes a preparatory stage: data preprocessing. This stage takes raw input data in the form of ontological individuals (structured knowledge fragments), and figures and transforms them into a format suitable for subsequent processing. This transformation occurs through various techniques, including text embedding, ontology embedding, and image embedding. These techniques essentially create numerical representations of the data, allowing the later stages of the FAM to efficiently exploit it. Figure 5 illustrates the architecture of the Filtering Article Module.

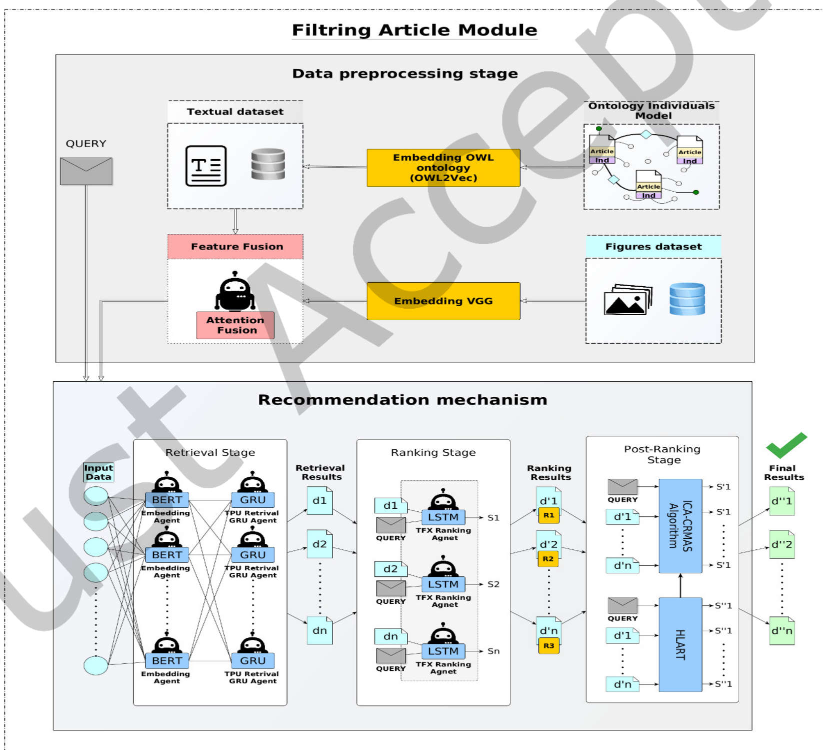


Fig. 5. The architecture of the Filtering Article Module

3.5.1 Preparing the data-set. Prior to feeding data into the FAM, we implement a novel data preprocessing pipeline to capture both textual and visual aspects of academic papers, maximizing the effectiveness of our recommendation system. This pipeline distinguishes our approach from previous work and paves the way for a unique contribution to citation recommendation. We leverage two distinct embedding techniques to capture the rich semantic information inherent in both textual and visual data. To represent the content of Onto-CCR, we employ the OWL2Vec technique, specifically tailored for ontologies. This technique transforms classes, properties, and individuals into vectors within a high-dimensional space, capturing their semantic relationships and contextual information. For visual components, we utilize a VGG convolutional neural network (CNN). This network extracts semantic features from figures, graphs, and diagrams, translating them into vectorized representations. By incorporating visual information, we enhance the comprehensiveness and accuracy of our retrieval process compared to solely text-based approaches. Once embeddings are generated, we fuse them into a unified representation using a deep learning-based attention mechanism. This method allows us to seamlessly integrate textual and visual data, capitalizing on their complementary strengths. Attention mechanisms dynamically assign weights to different input elements, prioritizing those most relevant to the query. By deploying this novel data preprocessing pipeline with a unique combination of OWL2Vec embedding and attention-based fusion, we provide a richer and more nuanced representation of academic papers, potentially enhancing the accuracy and effectiveness of our citation recommendation system.

3.5.2 Retrieval stage. The retrieval stage is vital in our system as it identifies relevant papers in response to user queries. We use a hybrid approach combining BERT and GRU models to determine similarity between queries and papers. This process efficiently navigates a large dataset, reducing search space and aiding focused exploration. Integrating BERT and GRU leverages their strengths for better retrieval performance. BERT enhances semantic comprehension, while GRU uncovers temporal nuances. This fusion results in more accurate and meaningful results for users, surpassing the capabilities of individual models.

3.5.3 Ranking stage. Enhancing relevance through LSTM-powered ranking involves using Long Short-Term Memory (LSTM) to meticulously assign scores based on parameters like paper quality, citation count, author reputation, and contextual relevance. LSTM's deep semantic analysis capability improves recommendation precision by generating a hierarchically ranked list of the most pertinent papers. Its versatility in modeling diverse data types enhances the ranking process, providing richer paper representations. LSTM's adaptive learning ensures resilience to evolving data patterns, sustaining effectiveness in paper ranking.

3.5.4 Post-Ranking Algorithm. The post-ranking stage, a crucial component of our FAM module, meticulously refines the ranked list of papers to deliver personalized and diversified recommendations. It builds upon the LSTM-ranked results, strategically employing HLART (Hierarchical Latent Aspect Rating Technique) and ICA-CRMAS algorithms to re-rank papers based on their latent aspects and user preferences. The ICA-CRMAS algorithm, grounded in ontology, orchestrates the final recommendation process within our FAM module. It commences by calculating word node power for each word within every article, utilizing three probabilities:

- PW: The probability of a word appearing in a scientific paper.
- PWC: The probability of a word appearing in a cited paper.
- PWA: The probability of a word appearing in other authors' papers.

Formula 2 encapsulates the computation of word node power:

$$Word\_Node\_Power = \sum_{i=0}^{m} Pw_i + \sum_{j=0}^{n} Pwc_j + \sum_{k=0}^{h} Pwa_k \tag{2}$$

Subsequently, article node power is calculated, incorporating three additional semantic relationships:

- CBP: The probability of Paper 1 citing Paper 2.
- PMA: The probability of mutual authors.

- CSP: The probability of both papers being cited in the same paper.

Formulas 3 delineate the computation of article node power:

$$Article\_Node\_Power_0 = \sum_{i=0}^{m} Cbp_i + \sum_{j=0}^{n} Pma_j + \sum_{k=0}^{h} Csp_k + \sum_{z=0}^{f} Word\_Node\_Power_{0_z} \quad (3)$$

$$Article\_Node\_Power_{id} = \left(\frac{P(Article\_Node\_Power_{id} \cap Article\_Node\_Power_{id-1})}{P(Article\_Node\_Power_{id-1})} \right) * \left(\sum_{i=0}^{m} Cbp_{id_i} + \sum_{j=0}^{n} Pma_{id_j} + \sum_{k=0}^{h} Csp_{id_k} + \sum_{z=0}^{f} Word\_Node\_Power_{id_z} \right), i > 0.$$

With calculated article powers, the algorithm proceeds to select the most relevant articles. Agents traverse ontology graph nodes, initiating from the highest-powered article and calculating subsequent article powers using conditional probabilities and a Markov chain framework. Bayes' theorem further refines the computations. This mechanism induces dynamic changes in article and word powers, leading to the elimination of nodes falling below a designated Learning Rate. Iterations witness a decrease in node powers and a corresponding reduction in article count until it aligns with user requirements. Entropy, employed to evaluate the contained information, increases with each iteration, reaching its maximum alongside the minimum article count, signaling the identification of the most relevant and representative set of articles that effectively address user needs (Refer to Appendix G for illustrative examples of The Algorithm Iterative Process in the Filtering Article Module).

4 Experiments and evaluations

4.1 Datasets and preprocessing

To assess the effectiveness of our proposed method, we employ two datasets: DBLP and Onto-CCR. These datasets are briefly introduced below, along with an outline of the preprocessing steps performed to enhance their suitability for the citation recommendation process. The DBLP dataset stands as a comprehensive repository of computer science publications, playing a crucial role in research discovery, scholarly collaboration, and AI development. Boasting over 6.9 million publications, detailed metadata, and a robust citation network, DBLP not only offers personalized citation recommendations but also serves as a benchmark for evaluating AI models and provides valuable insights into research trends. Encompassing contributions from 6,500 conferences and 1,800 journals, DBLP represents a diverse and expansive knowledge base.

The Onto-CCR dataset is a rich collection of scientific articles augmented with comprehensive citation information. This dataset serves as the primary source for training and testing our system, ensuring representation across diverse research areas. Acquisition of the dataset involved meticulous web scraping and crawling techniques, accompanied by measures such as source verification, data validation, and anomaly detection to ensure high data quality and credibility. What sets Onto-CCR apart from other citation recommendation datasets is its inclusion of images for each research paper, a novel feature introduced for the first time in citation recommendation datasets. Before embarking on the evaluation process, we meticulously preprocessed both the DBLP and Onto-CCR datasets to ensure their suitability for citation recommendation. To maintain data quality, we excluded papers with incomplete information or fewer than five citations. We further refined the content by filtering out words with two characters or less and applying stopword removal techniques. These measures aimed to distill the datasets down to their most meaningful and impactful words, resulting in a focused subset of 143,358 papers

Algorithm 1: ICA-CRMAS Post-Ranking Algorithm

```
Data: Pw, Pwc, Pwa, Cbp, Pma, Csp
input   : Articles_List, N_Words, Top_Nb_Articles
Output : Onto-CCR_Graph
Result: Recommended_Articles_List
Words_List ← Extract_Words(Articles_List);
Onto-CCR_Graph ← Initialization_Graph_Nodes(Words_List, Articles_List);
Initialize(Filtring_Agents , Recommended_Articles_Buffer);
begin
    Words_Probability_List ← Counting_Probability_Of_Existence(Words_List);
    Onto-CCR_Graph ← Update_Graph(Words_Probability_List);
    for Word ∈ Onto − CCR_Graph do
        Pw ← Assigning_Task_To_Agents(Word, Pw_Counting_Agents) ;
        Pwc ← Assigning_Task_To_Agents(Word, Pwa_Counting_Agents) ;
        Pwa ← Assigning_Task_To_Agents(Word, Pwc_Counting_Agents) ;
        Onto-CCR_Graph ← Update_Graph(Pw, Pwc, Pwa, Word);
    for Article ∈ Onto − CCR_Graph do
        Cbp ← Assigning_Task_To_Agents(Article, Cbp_Counting_Agents) ;
        Pma ← Assigning_Task_To_Agents(Article, Pma_Counting_Agents) ;
        Csp ← Assigning_Task_To_Agents(Article, Csp_Counting_Agents) ;
        Onto-CCR_Graph ← Update_Graph(Cbp, Pma, Csp, Article);
    while (Size(Articles_List) > Top_Nb_Articles  &&  Entropy_H < Learning_Rate_Entropy) do
        for (Node ∈ Onto − CCR_Graph) do
            if (Node ∈ Words_List ) then
                Assigning_Task_To_Agents(Node, Pw_Counting_Agents) ;
                Assigning_Task_To_Agents(Node, Pwa_Counting_Agents) ;
                Assigning_Task_To_Agents(Node, Pwc_Counting_Agents) ;
                    Word_Node_Power ← \sum_{i=0}^{m} Pw_i + \sum_{j=0}^{n} Pwc_j + \sum_{k=0}^{h} Pwa_k
                if (Word_Node_Power < Learning_Rate) then
                    Remove_Node_Agents ← Assigning_Task_To_Agents(Node) ;
                    Update_Graph_Agents ← Assigning_Task_To_Agents(Onto-CCR_Graph)
                else
                    Update_Graph_Agents ← Assigning_Task_To_Agents(Node, Word_Node_Power)
            else if (Node ∈ Articles_List) then
                Id ← GetArticleID(Node) ;
                Define_Predecessor_Agents ← Assigning_Task_To_Agents(Node, Onto-CCR_Graph);
                Assigning_Task_To_Agents(Node, Cbp_Counting_Agents) ;
                Assigning_Task_To_Agents(Node, Pma_Counting_Agents) ;
                Assigning_Task_To_Agents(Node, Csp_Counting_Agents) ;
                if (Id = 0 | Iteration = 1) then
                    Article_Node_Power_{id} ← \sum_{i=0}^{m} Cbp_i + \sum_{j=0}^{n} Pma_j + \sum_{k=0}^{h} Csp_k + \sum_{z=0}^{f} Word_Node_Power_{id_z}
                else if (Id > 0 & Iteration > 1) then
                    Article_Node_Power_{id} ← ( P(Article_Node_Power_{id} ∩ Article_Node_Power_{id-1}) / P(Article_Node_Power_{id-1}) ) *
                        ( \sum_{i=0}^{m} Cbp_{id_i} + \sum_{j=0}^{n} Pma_{id_j} + \sum_{k=0}^{h} Csp_{id_k} + \sum_{z=0}^{f} Word_Node_Power_{id_z} )
                if (Article_Node_Power_{id} < Learning_Rate) then
                    Remove_Node_Agents ← Assigning_Task_To_Agents(Node) ;
                    Update_Graph_Agents ← Assigning_Task_To_Agents(Onto-CCR_Graph) ;
                else
                    Update_Graph_Agents ← Assigning_Task_To_Agents(Node, Article_Node_Power_{id}) ;
        Update_Graph_Agents ← Assigning_Task_To_Agents(Articles_List)
            Entropy_H ← \sum p(Node_{id}) \log_b p(1 / Node_{id})
    Recommended_Articles_List ← Sorting_Descending_Order(Articles_List)
```

for DBLP and 85,341 papers for Onto-CCR.

The preprocessing of Onto-CCR followed a similar, yet more rigorous, approach due to its richer content. In addition to excluding low-quality papers, we implemented comprehensive noise reduction strategies. This involved eliminating special characters, punctuation, typos, and HTML tags that could potentially distort the analysis. Furthermore, we standardized the text by converting it to lowercase, removing whitespace inconsistencies, and applying consistent tokenization. To enhance the focus on relevant keywords, we systematically removed stopwords and employed linguistic techniques like stemming and lemmatization to reduce words to their base forms. These meticulous preprocessing steps significantly improve the accuracy and precision of the Onto-CCR dataset, preparing it for robust evaluation. The summary statistics of both datasets are presented in Table 2.

Datasets		Papers	Authors	Citations	Venues	Average citations
DBLP	Train	99,350	98,547	2,170,000	4413	5.78
	Test	44,008	42,234	930,000	1892	5.78
Onto-CCR	Train	62,403	192,730	3,539,760	4947	8.4
	Test	22,938	82,598	1,517,040	2121	8.4

Table 2. Statistics of used datasets.

4.2 Evaluation metrics

In evaluating the performance of our citation recommendation approach, we utilize several key metrics to assess the quality and utility of our recommendations. These metrics include Mean Reciprocal Rank (MRR), Recall at K, Precision at K (P@K), and factorized_top_k. MRR measures our ability to rank the most relevant citation highest in the list, indicating recommendation accuracy (Refer to Appendix H for detailed information regarding the formulas of the evaluation metrics).

4.3 Comparing with other approaches

To ensure a comprehensive and unbiased evaluation of our novel citation recommendation approach, we conducted a meticulous comparative analysis against a carefully selected set of prominent baseline models. These systems represent a diverse array of methodologies and recommendation formulations, collectively capturing the advancements and trends in context-aware citation recommendation. The rationale for their selection is twofold:

(1) Diversity of Approaches: This variety ensures a comprehensive assessment, as our model's performance is measured against distinct strategies for representing and recommending citations.
(2) State-of-the-Art Performance: Each baseline model has demonstrated superior performance in various evaluation metrics, including precision, recall, nDCG, MRR, and recommendation accuracy, within their respective domains. This establishes a high benchmark, ensuring a rigorous evaluation of our approach's effectiveness.

4.3.1 Brief Descriptions of Baseline Models.

- **TADW (Yang 2015)** [49]: Network Representation Learning with Rich Text Information. TADW incorporates text features into network representation learning using matrix factorization, outperforming other baselines in multi-class classification tasks, particularly in noisy and limited training data scenarios.
- **CCF (LIU 2015)[37]**: Context-Based Collaborative Filtering for Citation Recommendation. CCF employs association mining on citation relations to recommend papers, considering co-occurring citing papers as similar. It has shown significant improvements in precision, recall, and recommendation metrics.

- **PCCF (Sugiyama 2015)** [46]: Citation Recommendation with Adaptive Neighbor Selection. Focused on recommendation accuracy, PCCF employs adaptive neighbor selection, demonstrating superior performance in terms of nDCG and MRR.
- **Paper2vec (Ganguly 2017)**[26]: Combining Graph and Text for Scientific Paper Representation. Paper2vec creates rich paper representations by combining textual and graph-based information, outperforming state-of-the-art methods in node classification and link prediction tasks.
- **NN-CCR (Ebesu 2017)**[22]: Neural Citation Network for Context-Aware Citation Recommendation. NN-CCR uses an encoder-decoder architecture to recommend citations based on the citation context, significantly improving over competitive baselines in terms of recommendation accuracy.
- **CNCRec (Wang 2022)**[72]: Collaborative Filtering with Network Representation Learning for Citation Recommendation. CNCRec combines network representation learning and collaborative filtering, outperforming state-of-the-art methods in precision, recall, and MRR, addressing challenges such as data sparsity.
- **DL-CCR (Abbas 2022)** [55]. Leveraging deep learning, this paper addresses the cold-start problem in citation recommendation by classifying research articles into predefined linguistic categories and computing similarity using rhetorical zone embeddings. Experimental results showcase a 76% accuracy in rhetorical zone classification and an average nDCG score of 0.704 in ranked recommendations.
- **ECCR (Dinh 2023)** [19]. Introducing a novel approach to automated citation recommendation systems, ECCR leverages Bidirectional Encoder Representations from Transformers (BERT) and a deep neural auto-encoding mechanism with a self-attention-based module. Empirical evaluations demonstrate superior performance on the arXiv dataset.
- **CRB-HDGCN (DAI 2023)** [17]. CRB-HDGCN enhances BERT's understanding of citations through title-enriched training and graph network analysis, delivering significantly improved citation recommendations compared to existing methods.

4.3.2 Experiments with DBLP dataset. To assess the competitiveness of our ICA-CRMAS approach, we conducted comparative analyses with state-of-the-art models on the DBLP dataset. The results, captured in Table 3, paint a compelling picture of ICA-CRMAS's exceptional capabilities. At the forefront of performance, ICA-CRMAS stands as an exemplar of exceptional performance. The system showcases unparalleled accuracy, particularly evident in its Precision@5 metric, which attains a noteworthy 55.07%. This achievement positions ICA-CRMAS as a leader, surpassing its closest competitor, CRB-HDGCN DAI(2023), by a significant margin of 7.53 percentage points. To contextualize this advantage, it implies that, out of every 100 recommendations, 7.4 exhibit superior precision when facilitated by ICA-CRMAS. This precision superiority extends further, as evidenced by a substantial lead over Abbas(2022) and CNCRec(Wang 2022), standing at 7.7% and 15.75%, respectively. Across all N values (5, 10, 20), ICA-CRMAS maintains consistent dominance in precision, solidifying its place as the superior approach. ICA-CRMAS not only excels at pinpointing top relevant citations but also casts a wider net to capture valuable connections. Its Recall@5 of 35.02% outshines all other models, exceeding CRB-HDGCN DAI(2023) by 5.87% points. This signifies ICA-CRMAS's ability to retrieve a larger proportion of relevant items within the top 5 recommendations. The lead over Abbas(2022) and CNCRec(Wang 2022) is even more pronounced, with advantages of 11.57% and 15.22% points, respectively. This dominance persists across all N values, demonstrating ICA-CRMAS's comprehensive recall capabilities.

To understand how efficiently ICA-CRMAS ranks relevant citations, we turn to Mean Reciprocal Rank (MRR). At MRR@5, ICA-CRMAS scores a remarkable 64.61%, outperforming CRB-HDGCN DAI(2023) by 6.07% points. This indicates that relevant citations consistently rank higher in ICA-CRMAS ommendations, saving researchers valuable time and effort in navigating the literature. The lead over Abbas(2022) and CNCRec(Wang 2022) is equally impressive, with advantages of 13.06% and 17.29% points, respectively. This consistent dominance across N values underscores ICA-CRMAS's exceptional efficiency in prioritizing relevant citations. We observe that TADW (Yang 2015) [49] performs noticeably inferiorly due to its reliance on word frequency compared to other methods. CCF (LIU 2015) [37] significantly outperforms TADW (Yang 2015), underscoring the efficacy of incorporating additional features, such as citation relations, as source data. This suggests that the inclusion of more diverse features is considerably more beneficial than relying solely on word frequency to identify pertinent papers. By incorporating author profile features, PCCF (Sugiyama 2015) [46] further enhances performance, highlighting the pivotal role played by author-related features in the citation recommendation task. Both Paper2vec (Ganguly 2017) and CNCRec (Wang 2022) leverage an embedding approach, utilizing both textual and graph-based information. This methodology demonstrates its superiority over the aforementioned methods, which are based solely on either text representation or graph-based information.

The models NCN (Ebesu 2017), Abbas (2022) [55], and CCR (Dinh 2023) [19] leverage latent information through a Context-aware approach. This underscores the importance of integrating auxiliary information to aid in the citation contextual learning process, providing a wealth of valuable semantic information that contributes to delivering high-quality candidates within the context of citation recommendation. CRB-HDGCN (DAI 2023) [17] notably improves performance compared to all aforementioned models. By combining citation-enhanced pretrained language models (PLM) with deep heterogeneous graph learning, this research yields competitive results vis-a-vis our system. This stems from its ability to harness the advantages of several mentioned models, such as merging graph-based information with text representation, utilizing contextual information, and relying on PLM models. However, it has some shortcomings. It overlooks certain contextual features in the context of citation recommendation, stemming from a lack of reliance on a robust structural knowledge base. It does not employ post-filtering or semantic rules to refine final outputs. Moreover, it neglects the application of multimodal techniques that consider both textual and image features for recommending relevant papers. This justifies the superiority of our ICA-CRMAS system.

Our system not only relies on a robust contextual representation of the citation recommendation context and employs deep learning models but also adopts multimodal techniques, considering both textual and visual elements. Furthermore, it applies filtering at all stages of recommendation—pre-filtering within the model and post-filtering through a unique innovative algorithm based on a multi-agent system. Ultimately, our system exhibits high transparency compared to other models. It can display all semantic rules used to generate relevant papers, providing a clear vision to the user and increasing its trustworthiness.

4.3.3 Efficiency evaluation. ICA-CRMAS has consistently demonstrated its commitment to retrieval accuracy, as evidenced by consistently achieving the highest R@20 scores across all sample sizes (Table 6). This unwavering focus on accuracy aligns with the primary objective of addressing users' information needs, ensuring that recommendations provided are not only timely but also directly relevant to their research queries. In contrast, despite ECCR's ability to offer faster test times in specific scenarios, the noticeable decline in its R@20 scores highlights a consequential trade-off that emphasizes speed at the expense of retrieval relevance. In this context, ICA-CRMAS stands out by successfully striking a more advantageous balance between efficiency and accuracy. ICA-CRMAS demonstrates a remarkable ability to balance efficiency and accuracy, as evidenced by its consistently superior R@20 scores (retrieval accuracy) across all sample sizes in Table 6. Notably, it maintains highly competitive test times, indicating that highly relevant recommendations are generated promptly. This balance is particularly commendable, as it ensures user satisfaction without sacrificing timely responses. While ICA-CRMAS exhibits slightly longer training times than some competitors, the investment is readily justified by the substantial gains

Methods	Precision@N			Recall@N			MRR@N		
	5	10	20	5	10	20	5	10	20
TADW (Yang 2015) [49]	0.3087	0.2911	0.2438	0.1581	0.1812	0.2231	0.4622	0.4621	0.4937
CCF (LIU 2015) [37]	0.3321	0.3112	0.2701	0.1678	0.2077	0.2468	0.4828	0.4921	0.5018
PCCF (Sugiyama 2015) [46]	0.3217	0.2983	0.2508	0.1638	0.1837	0.2235	0.4781	0.4838	0.5167
Paper2vec (Ganguly 2017) [26]	0.3523	0.3402	0.2843	0.1728	0.1883	0.2428	0.5012	0.5123	0.5189
NCN (Ebesu 2017) [22]	0.3462	0.3245	0.2783	0.1783	0.2051	0.2443	0.4932	0.5034	0.5541
CNCRec (Wang 2022) [72]	0.3831	0.3525	0.3012	0.1932	0.2168	0.2541	0.5123	0.5234	0.5538
Abbas (2022) [55]	0.4015	0.3555	0.3115	0.2005	0.2255	0.2655	0.5255	0.5425	0.5755
ECCR (Dinh 2023) [19]	0.4259	0.3796	0.3412	0.2297	0.2467	0.2895	0.5547	0.5656	0.6046
CRB-HDGCN (DAI 2023) [17]	0.4754	0.4395	0.3954	0.2815	0.3054	0.3452	0.5954	0.6054	0.6445
ICA-CRMAS	0.5507	0.5254	0.4299	0.3502	0.3695	0.3747	0.6461	0.6497	0.6530

Table 3. Recommendation performance comparisons on DBLP datasets in terms of PrecisionN, RecallN, and MRRN.

in accuracy: the most crucial metric for citation recommendation systems. This positive correlation between training time and retrieval accuracy underscores ICA-CRMAS's commitment to prioritizing highly relevant recommendations, ultimately enhancing user research productivity and knowledge discovery.

While acknowledging ICA-CRMAS's strengths, the observed variations in training times present a fertile ground for future research aimed at further optimizing the system's efficiency. Future studies could delve into enhancing ICA-CRMAS's training efficiency to solidify its position as a leading solution in the domain of citation recommendation systems. It is imperative to recognize that the ideal balance between accuracy and efficiency may vary across different application contexts and user preferences. Consequently, future investigations could explore mechanisms for tailoring ICA-CRMAS to align with diverse usage scenarios, thereby offering a personalized and adaptable solution to users with varying needs. In conclusion, for users who prioritize the most relevant and accurate citation suggestions, ICA-CRMAS stands out as the superior choice. Its consistent focus on accuracy, coupled with a balanced approach to efficiency, positions it as a compelling solution for real-world applications where the relevance of recommendations is of paramount importance. By consistently delivering the most pertinent citations to users, ICA-CRMAS holds significant potential to enhance research productivity and knowledge discovery, contributing to advancements across various academic fields.

Fig. 6. Efficiency evaluation on DBLP.

Methods	10% samples			50% samples			100% samples		
	Train	Test	R@20	Train	Test	R@20	Train	Test	R@20
ECCR (Dinh 2023)	35,544 s	2,544 s	0.1297	185,635 s	11,435 s	0.2467	354,754 s	22,980 s	0.2895
CRB-HDGCN (DAI 2023)	30,466 s	3,467 s	0.1815	154,354 s	17,954 s	0.2854	308,467 s	35,908 s	0.3452
ICA-CRMAS	37,405 s	4,295 s	0.1611	188,054 s	15,545 s	0.3037	345,665 s	32,092 s	0.3747

4.4 Experiments with Onto-CCR dataset dataset

In our rigorous evaluation, we compared ICA-CRMAS with two prominent citation recommendation models: ECCR (Dinh 2023) and CRB-HDGCN (DAI 2023), utilizing the unique Onto-CCR dataset. This dataset is distinctive

Model	Top 1	Top 5	Top 10	Top 50	Top 100
CRB-HDGCN	0.0070 (0.70%)	0.1500 (15%)	0.2800 (28%)	0.5357 (53.57%)	0.7800 (78%)
ECCR	0.0085 (0.85%)	0.1850 (18.5%)	0.3200 (32%)	0.5892 (58.92%)	0.8200 (82%)
ICA-CRMAS	0.0105 (1%)	0.2104 (21%)	0.3605 (36%)	0.6326 (63%)	0.8707 (87%)

Table 4. Recommendation performance comparisons on Onto-CCR datasets in terms of factorized_top_k.

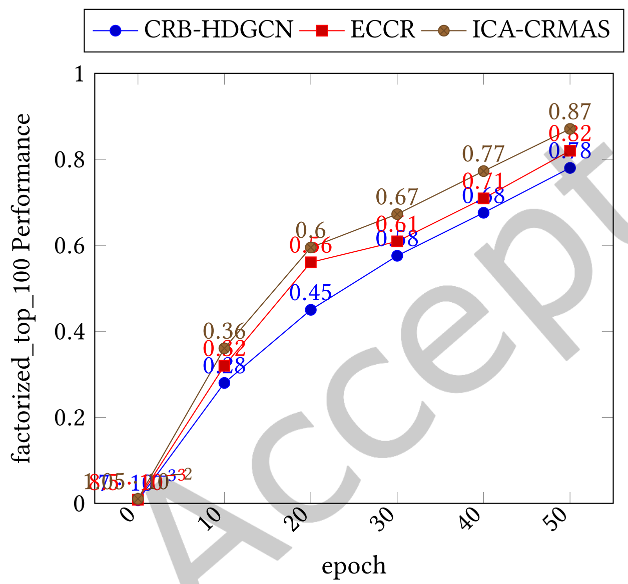


Fig. 7. Curve of performance for 'factorized_top_100' over 50 epochs.

for combining text-based features and images linked to citations, offering a comprehensive context for evaluating model effectiveness. ICA-CRMAS demonstrated superior performance across all top-k accuracy metrics. Achieving a stellar 1.05% at Top 1, it outperformed both ECCR (0.85%) and CRB-HDGCN (0.70%) (Figure 5). This trend continued, with ICA-CRMAS leading at Top 5 (21.04%), Top 10 (36.05%), Top 50 (63.26%), and Top 100 (87.07%) accuracy (Figure 7), emphasizing its capability to provide accurate recommendations across various scenarios (Table 5).

The key to ICA-CRMAS's success lies in its adept use of both textual and visual information. Unlike competitors focusing solely on text analysis, ICA-CRMAS leverages sophisticated image embedding techniques. This allows it to delve beyond surface-level textual relationships, uncovering intricate connections between citations and associated visual content. This enriched contextual understanding empowers ICA-CRMAS to recommend citations aligning not just with textual information but also seamlessly with visual cues, especially valuable in domains where images and diagrams play a crucial role in knowledge representation.

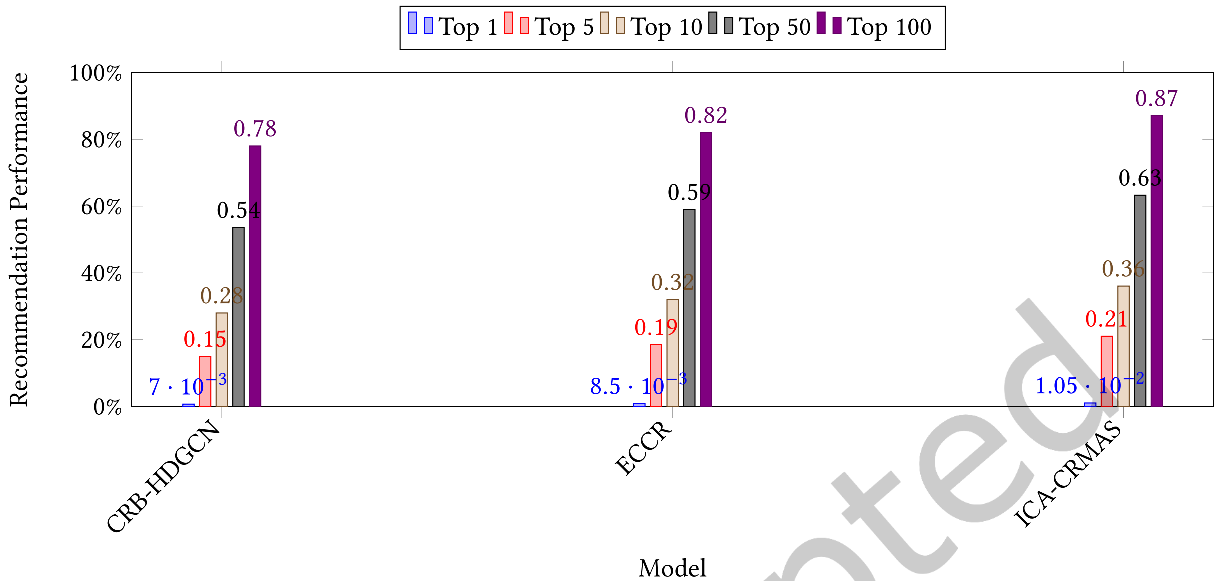


Table 5. Recommendation performance comparisons on Onto-CCR datasets in terms of factorized_top_k.

4.5 Excision Study on ICA-CRMAS Components

In this section, we conduct an excision study on various variants of the Intelligent Context-Aware Citation Recommendation System (ICA-CRMAS), focusing on their performance across different configurations. These variants are systematically tested using textual data only, images only, and a combined approach incorporating both text and images. The evaluation criteria include essential metrics such as Precision@N, Recall@N, and Mean Reciprocal Rank (MRR), assessed at N values of 5, 10, and 20. Through this rigorous analysis, we aim to discern the impact of different modules within ICA-CRMAS on citation recommendation accuracy and relevance. Furthermore, this excision study encompasses an investigation into computational costs and overall accuracy metrics associated with each variant. The table 6 showcases the variants of ICA-CRMAS along with their respective configurations.

Group	**Variant Name**	**Text data**	**Image data**	**FAM**	**SAM & QMM**	**ICM & OM**
1	FAM-Text-SAM-QMM	✓		✓	✓	
	FAM-Img-SAM-QMM		✓	✓	✓	
	FAM-Text-Img-SAM-QMM	✓	✓	✓	✓	
2	FAM-Text-ICM	✓		✓		✓
	FAM-Img-ICM		✓	✓		✓
	FAM-Text-Img-ICM	✓	✓	✓		✓
3	ICA-CRMAS-Tex	✓		✓	✓	✓
	ICA-CRMAS-Img		✓	✓	✓	✓
	ICA-CRMAS-Text-Img	✓	✓	✓	✓	✓

Table 6. Summary of ICA-CRMAS Variants

We designed nine configurations for the ICA-CRMAS system, categorized into three main groups based on their functionalities:

- Group 1 : This group leverages three core modules: Filtering Module (FAM), Searching Module, and Query Modeling Module. The Filtering Module plays a critical role, training on data and generating final predictions. The Searching and Query Modeling Modules function as a single unit due to their interdependence and collaborative role. This group will be tested with three different input types: text only, image only, and a combination of text and image inputs. Specific models in this group include FAM-Text-SAM-QMM, FAM-Img-SAM-QMM, and FAM-Text-Img-SAM-QMM.
- Group 2 : This group focuses on configurations that integrate Filtering Module, Intelligent Context Management (ICM) and Ontology Module (OM). Here, the Searching and Query Modeling Modules are excluded. This is due to the synergistic relationship between ICM and Ontology Module. The ICM leverage the semantic representations generated by Ontology Module for enhanced performance. Similar to Group 1, this group will undergo testing with three input scenarios: text only, image only, and a combination of text and image inputs. Models belonging to this group are FAM-Text-ICM, FAM-Img-ICM, and FAM-Text-Img-ICM.
- Group 3 : The third group encompasses configurations that utilize all components of ICA-CRMAS. These models will be evaluated using the same three input variations as the previous groups: text only, image only, and both text and image inputs. Models in this group include ICA-CRMAS-Tex, ICA-CRMAS-Img, and ICA-CRMAS-Text-Img.

It's important to note that the ICA-CRMAS variants utilizing image data incorporate the VGG convolutional neural network (CNN) model within their architecture. This CNN module is specifically designed to handle and extract features from image data. Conversely, the variants that solely rely on text data do not include the VGG CNN module, as their focus lies on processing and analyzing textual information.

4.5.1 Evaluating ICA-CRMAS Variant Performance. To assess the effectiveness of various ICA-CRMAS configurations, we conducted comparative analyses on the DBLP dataset (Table 2). Table 7 summarizes the performance results obtained from experiments with the nine models. The results clearly demonstrate the advantage of utilizing multimodal inputs (text and image) within the ICA-CRMAS framework. The model leveraging both text and image data (FAM-Text-Img-SAM-QMM and ICA-CRMAS-Text-Img) consistently outperforms models that rely solely on text or image inputs across all three evaluation metrics (Precision@N, Recall@N, and MRR@N).

- Precision@N: The multimodal models achieve a significant improvement in Precision@N compared to text-only models (FAM-Text-SAM-QMM). At N=20, the improvement is around 13% (0.4250 vs. 0.3850). This indicates that the inclusion of image data alongside text data leads to a higher proportion of relevant retrievals among the top 20 results.
- Recall@N: Similar trends are observed for Recall@N. The multimodal models achieve a better retrieval rate of relevant information compared to text-only or image-only models. At N=20, the improvement for the best multimodal model (ICA-CRMAS-Text-Img) is approximately 24% (0.3747 vs. 0.3050) compared to the text-only model.
- MRR@N: The Mean Reciprocal Rank (MRR@N) also reflects the superiority of multimodal models. They achieve a higher average ranking position for relevant retrievals within the top N results. For instance, at N=20, the best multimodal model (ICA-CRMAS-Text-Img) achieves an MRR@N of 0.6530, which is a significant improvement of 24% compared to the text-only model (0.5291).

It's important to mention that the image-only models (FAM-Img-SAM-QMM, FAM-Img-ICM, ICA-CRMAS-Img) exhibit considerably lower performance across all metrics compared to text-only or multimodal models. This suggests that relying solely on image data for scholarly paper recommendation tasks might not be sufficient.

Methods	Precision@N			Recall@N			MRR@N		
	5	10	20	5	10	20	5	10	20
FAM-Text-SAM-QMM	0.4050	0.3900	0.3750	0.2750	0.2900	0.3050	0.5250	0.5300	0.5350
FAM-Img-SAM-QMM	0.1592	0.1536	0.1478	0.1254	0.1302	0.1341	0.2921	0.2952	0.2973
FAM-Text-Img-SAM-QMM	0.4550	0.4400	0.4250	0.3250	0.3400	0.3550	0.6050	0.6064	0.6159
FAM-Text-ICM	0.4100	0.3950	0.3800	0.2800	0.2950	0.3100	0.5300	0.5350	0.5400
FAM-Img-ICM	0.0565	0.0555	0.0545	0.0458	0.0470	0.0482	0.1825	0.1842	0.1885
FAM-Text-Img-ICM	0.4600	0.4450	0.4300	0.3300	0.3450	0.3600	0.6100	0.6177	0.6200
ICA-CRMAS-Text	0.4250	0.4076	0.3850	0.2850	0.3089	0.3150	0.6250	0.6291	0.6350
ICA-CRMAS-Img	0.2057	0.1954	0.1749	0.1552	0.1595	0.1697	0.3050	0.3150	0.3201
ICA-CRMAS-Text-Img	0.5507	0.5254	0.4299	0.3502	0.3695	0.3747	0.6461	0.6497	0.6530

Table 7. Performance Comparison of ICA-CRMAS Variants on DBLP Dataset using PrecisionN, RecallN, and MRRN.

However, the improvements observed in the multimodal models compared to text-only models indicate that image data can be a valuable source of complementary information, enhancing the overall retrieval effectiveness.

The line chart (Figure 8) showcases the training performance of various ICA-CRMAS configurations on the Onto-CCR dataset. As the training progresses (x-axis), all models exhibit an upward trend in accuracy (y-axis), indicating successful learning. Notably, models that leverage both text and image data (FAM-Text-Img-SAM-QMM, FAM-Text-Img-ICM, ICA-CRMAS-Text-Img) consistently achieve up to 10% higher accuracy at epoch 50 compared to those relying solely on text. This quantifies the clear advantage of incorporating multimodal information for improved performance.

For instance, FAM-Text-Img-SAM-QMM, a multimodal model, achieves an accuracy of 0.82 at epoch 50, surpassing the 0.75 accuracy of the text-only model FAM-Text-SAM-QMM. This signifies a 7.0% improvement. Similar improvements are observed for FAM-Text-Img-ICM (4.0%) and ICA-CRMAS-Text-Img (10%) compared to their text-only counterparts. Conversely, models that solely rely on image data (FAM-Img-SAM-QMM, FAM-Img-ICM, ICA-CRMAS-Img) exhibit the lowest overall accuracy across all epochs, reaching a maximum of around 0.3 at epoch 50. This suggests that image data alone might be insufficient for accurate paper recommendation.

The loss curves (Figure 9) complement these findings. As expected, all curves show a generally downward trend throughout training, indicating that the models are learning and reducing prediction errors (loss). Here, multimodal models consistently achieve lower loss values compared to single modality models, reinforcing their superior performance. For example, ICA-CRMAS-Text-Img has a significantly lower loss (around 0.85) at epoch 50 compared to ICA-CRMAS-Text (around 1.89). However, this benefit comes at the cost of increased training times (Table 8, Figure 10).

Our analysis reveals no significant differences between the performance of models using the Group 1 approach (SAM and QMM) and the Group 2 approach (ICM and OM). Both models perform closely, with the FAM-Text-SAM-QMM model achieving an accuracy of 0.75 and the FAM-Text-ICM model achieving 0.79, a difference of only 0.04. This convergence in performance becomes even more evident with the FAM-Text-Img-SAM-QMM (0.82 accuracy) and FAM-Text-Img-ICM (0.83 accuracy) models, where the difference narrows to just 0.01. However, the Group 1 models based on SAM and QMM (particularly FAM-Img-SAM-QMM) achieve notably better results when using image features alone compared to the FAM-Img-ICM model. This improvement likely stems from the ability of FAM-Img-SAM-QMM to capture missing data, especially in the image domain, offering an advantage over the FAM-Img-ICM model. Overall, our observations reveals that the ICM, SAM, and QMM methods perform best

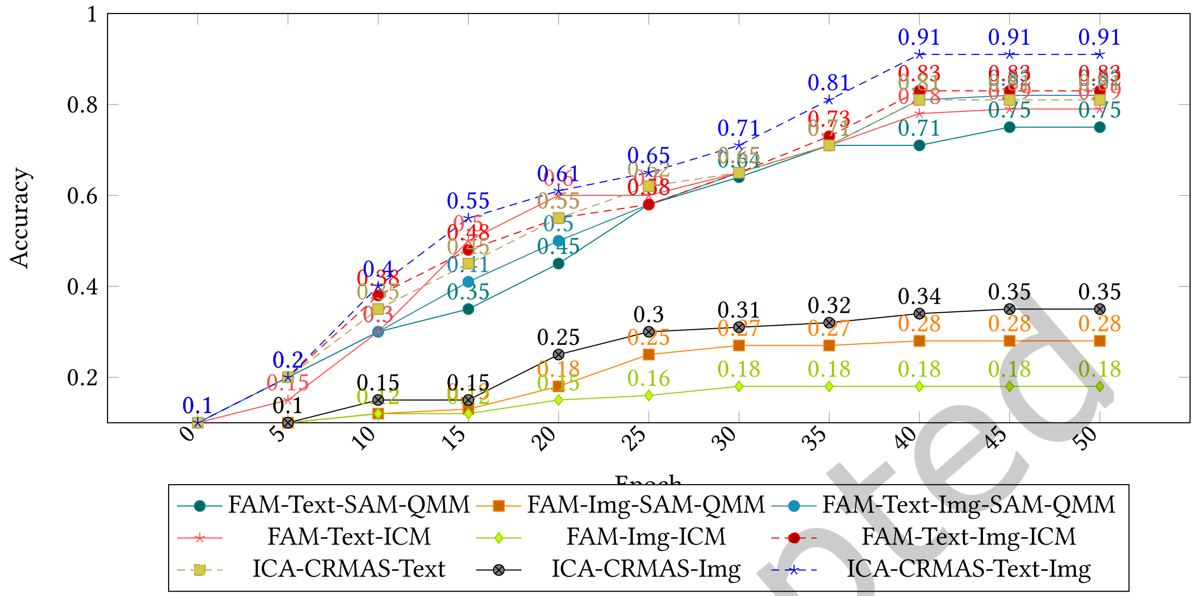


Fig. 8. Training accuracy comparison of ICA-CRMAS variants on Onto-CCR dataset

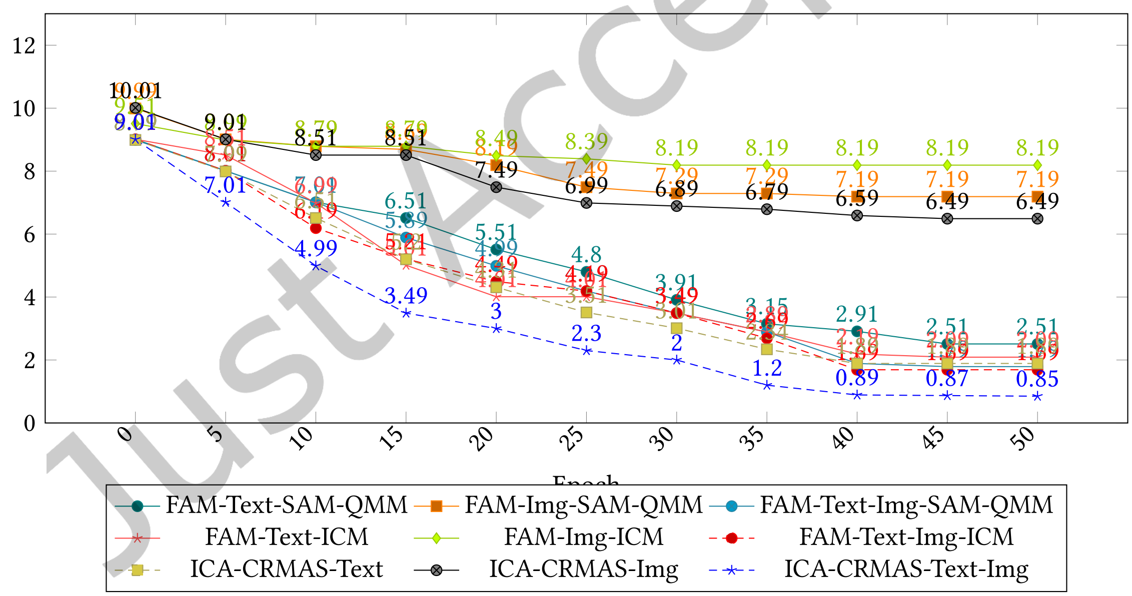


Fig. 9. Loss comparison of ICA-CRMAS variants on Onto-CCR dataset

Table 8. Computation cost in time for each ICA-CRMAS Variants.

Methods	Train	Test	Accuracy	Loss
FAM-Text-SAM-QMM	294,258 s	21,897 s	0.75	2.51
FAM-Img-SAM-QMM	76,087 s	7,063 s	0.28	7.19
FAM-Text-Img-SAM-QMM	328,382 s	30,487 s	0.82	1.79
FAM-Text-ICM	264,832 s	19,707 s	0.79	2.09
FAM-Img-ICM	72,082 s	6,691 s	0.18	8.19
FAM-Text-Img-ICM	311,098 s	28,882 s	0.83	1.69
ICA-CRMAS-Text	294,258 s	21,897 s	0.81	1.89
ICA-CRMAS-Img	80,092s	7,435 s	0.35	6.49
ICA-CRMAS-Text-Img	345,665 s	32,092 s	0.91	0.85

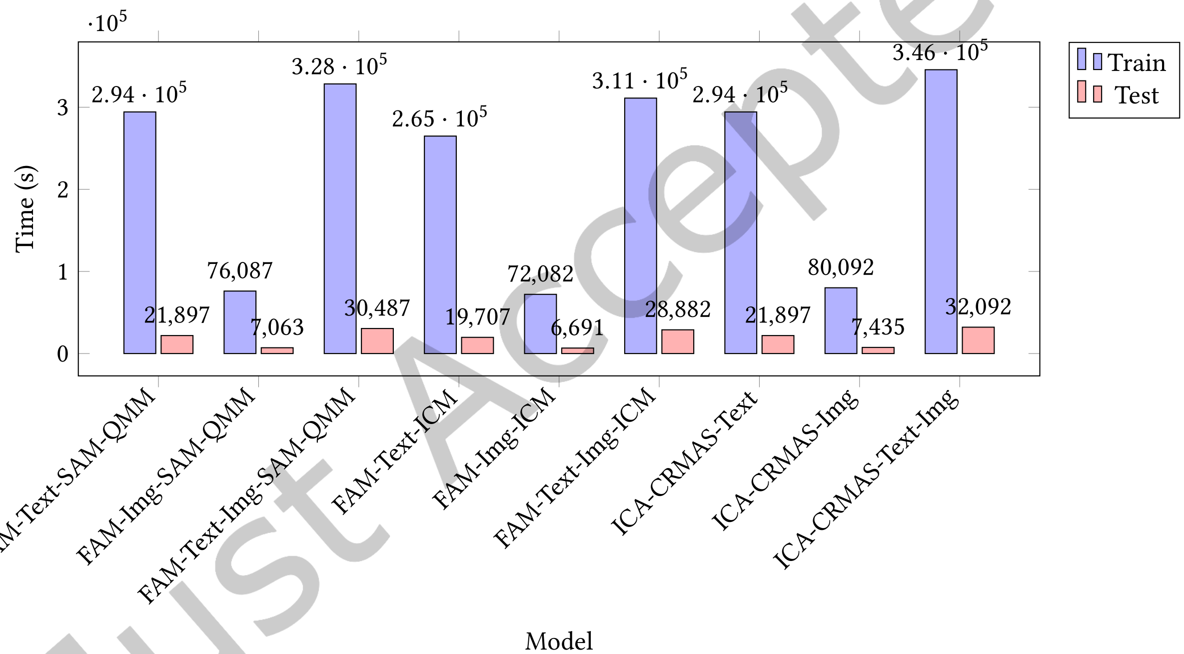


Fig. 10. Comparison of Training and Testing Time Across the ICA-CRMAS variants

when combined for processing both text and image data (0.91 accuracy), or when applied solely to text data (0.81 accuracy). Conversely, their performance appears to be lower when limited to images data only (0.35 accuracy). In conclusion, both accuracy and loss curves emphasize the benefits of incorporating multimodal data (text and image) within the ICA-CRMAS system(Figure 8, 9). These models achieve demonstrably superior performance with significant improvements in accuracy and lower loss values compared to single modality models.

5 System Usability Evaluation

To assess the usability of the ICA-CRMAS system, a user study was conducted utilizing the System Usability Scale (SUS) [6]. The SUS questionnaire comprises ten questions, each focusing on a specific aspect of the system's usability:

(1) **Frequent Use (FU)**: I would use this system often.
(2) **Ease of Use (EU)**: The system was easy to use.
(3) **System Functions Integration (FI)**: The system's functions worked well together.
(4) **Learning Curve (LC)**: Most people could learn to use the system quickly.
(5) **Confidence in the System (CO)**: I felt confident using the system.
(6) **System Complexity (SC)**: The system was unnecessarily complex.
(7) **Need for Support (NS)**: I would need technical support to use the system.
(8) **System Inconsistencies (SI)**: The system was inconsistent.
(9) **Cumbersome System (CU)**: The system was awkward to use.
(10) **Need for Training (NT)**: I needed a lot of training to use the system.

A total of 187 participants (96 male, 91 female) were recruited from research groups across various disciplines. Table 9 provides a detailed overview of the participant demographics (Refer to Appendix I for Detailed Participant Group Information). Participants were introduced to the ICA-CRMAS system and given sufficient time to explore its features independently. They were then asked to complete the SUS questionnaire, providing their honest feedback on the system's usability (Refer to Appendix J for Detailed procedures employed in our user study).

Table 9. Participant Demographics

Characteristic	**Category**	**Frequency**
Gender	Male	96
	Female	91
Research Role	Researcher	115
	Student	72
Years of Experience	0-5 years	85
	6-10 years	63
	11+ years	39

5.1 Brief Descriptions of Baseline Platforms

Google Scholar

Google Scholar functions as a web-based academic search engine and citation database, indexing over 170 million scholarly articles (as of June 2023) and handling more than 3 billion daily search queries [3]. Widely known for its massive database, user-friendly interface, and integrated citation recommendations, Google Scholar is, however, challenged by limited context awareness in its recommendations, heavy reliance on textual features, and the potential for overwhelming information density.

Base-Search

Base-Search serves as a citation recommendation system and academic search engine with a semantic analysis focus. It indexes over 85 million scholarly articles (as of June 2023) and has a growing user base of over 1 million registered researchers [2]. Noteworthy for its context-aware recommendations, semantic search capabilities, personalized suggestions, and a clean interface, Base-Search faces challenges such as a smaller database compared to Google Scholar, lower user adoption, and the risk of generating overly specific recommendations.

5.2 Results and Comparisons

The ICA-CRMAS system achieved an impressive Overall SUS Score of **76.73**, indicating a very positive user experience. This score compares favorably to other citation recommendation systems (Table 12). The SUS questionnaire categorizes questions into Positives (odd-numbered) and Negatives (even-numbered). Higher scores (1-5) in Positives reflect strengths, while for Negatives, they highlight areas for improvement.

ICA-CRMAS stands out against popular systems like Google Scholar and Base-Search (Tables 10 and 11), excelling in frequent use, seamless integration, and user confidence (Figure 11). While ICA-CRMAS performs well, Google Scholar had slightly higher scores in specific areas like ease of use and learning curve, revealing potential avenues for enhancement. Base-Search demonstrates noteworthy results, highlighting robust competition within the field. To gain a comprehensive picture of usability, we calculate the Overall SUS Score using established methods [6] (Algorithm 3). This score, incorporating both strengths and weaknesses, provides a valuable benchmark for system comparison. By analyzing detailed item-level scores and the Overall SUS Score, we can refine ICA-CRMAS to further enhance its user experience and solidify its position as a leading citation recommendation system.

$$
\begin{aligned}
\text{Adjusted Score}_{\text{odd}} &= (2.5 \times (\text{Original Score}_{\text{odd}} - 1)) \\
\text{Adjusted Score}_{\text{even}} &= (2.5 \times (5 - \text{Original Score}_{\text{even}})) \\
\text{Overall SUS Score} &= \frac{1}{10} \times \sum_{i=1}^{10} \text{Adjusted Score}_i \\
&= \frac{1}{10} \times ((\text{FU}_{\text{adjusted}} + \text{EU}_{\text{adjusted}} + \text{FI}_{\text{adjusted}} + \text{LC}_{\text{adjusted}} + \text{CO}_{\text{adjusted}}) \\
&\quad + (5 - \text{SC}_{\text{original}}) + (5 - \text{NS}_{\text{original}}) + (5 - \text{SI}_{\text{original}}) \\
&\quad + (5 - \text{CU}_{\text{original}}) + (5 - \text{NT}_{\text{original}}))
\end{aligned}
$$

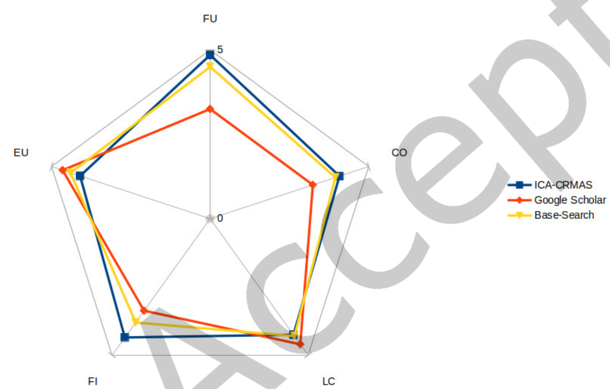
ICA-CRMAS emerges as a usability frontrunner, securing an impressive overall SUS score of 76.73, surpassing Google Scholar (63.75) and Base-Search (74.565). This achievement highlights its commitment to user-centric design, prioritizing ease of use, intuitive navigation, and a seamless experience. While Google Scholar grapples with challenges related to limited context awareness in recommendations, a significant reliance on textual features, and the potential for overwhelming information density. Similarly, Base-Search faces hurdles, including lower user adoption and the potential risk of generating overly specific recommendations.

Notably, ICA-CRMAS excels in positive SUS items, boasting a score of 54.24, indicating a user-friendly and empowering system. However, room for improvement remains, particularly in negative SUS items, where ICA-CRMAS scored 49.125. Users highlighted aspects needing refinement. Addressing these shortcomings could further elevate ICA-CRMAS's usability and boost its overall SUS score.

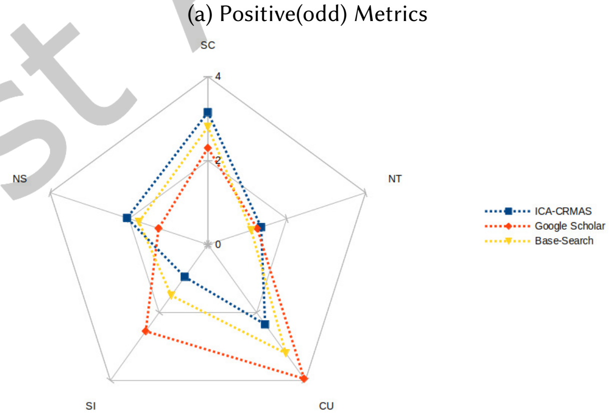
Despite these identified areas for improvement, ICA-CRMAS stands out as a compelling choice for researchers. Its intuitive design, seamless integration, and confidence-inspiring performance position it favorably in the research landscape. Ongoing attention to user-centric design and the incorporation of feedback will ensure the sustained excellence of ICA-CRMAS, enabling it to cater even more effectively to the evolving needs of researchers.

	ICA-CRMAS	Google Scholar	Base-Search
FU (Frequent Use)	4.85	3.25	4.50
EU (Ease of Use)	4.10	4.65	4.40
FI (System's Functions Integration)	4.35	3.37	3.80
LC (Learning Curve)	4.25	4.60	4.30
CO (Confidence in the System)	4.08	3.24	3.95

Table 10. Comparing Positive(odd) Usability Metrics between ICA-CRMAS, Google Scholar, and Base-Search with Adjusted Values



(a) Positive(odd) Metrics



(b) Negative(even) Metrics

Fig. 11. Comparing Negative(even) & Negative(even) Usability Metrics

	ICA-CRMAS	Google Scholar	Base-Search
SC (System Complexity)	3.15	2.30	2.80
NS (Need for Support)	2.05	1.25	1.75
SI (System Inconsistencies)	0.95	2.55	1.50
CU (Cumbersome System)	2.35	3.95	2.70
NT (Need for Training)	1.35	1.25	1.10

Table 11. Comparing Negative(even) Usability Metrics between ICA-CRMAS, Google Scholar, and Base-Search with Adjusted Values

	ICA-CRMAS	Google Scholar	Base-Search
Positive SUS Items (Adjusted)	54.24	45.84	49.88
Negative SUS Items (Adjusted)	49.125	61.65	49.25
Overall SUS Score	**76.73**	**63.75**	**74.565**

Table 12. Overall SUS Scores

6 Conclusions

In summary, as the volume of scientific information grows, ICA-CRMAS serves as a reliable guide for researchers in the vast sea of data. This Intelligent Context-based Recommender System tackles challenges such as information overload and inefficient paper recommendations, providing researchers with enhanced efficiency and accuracy. Its strength lies in a deep understanding of the research context. Going beyond precision, ICA-CRMAS utilizes advanced deep learning and multimodal learning, analyzing not only text but also figures and visual elements in research papers. This approach uncovers hidden connections, resulting in tailored recommendations that align with specific research interests and goals. Ensuring contextual relevance, ICA-CRMAS incorporates contextual information directly within its recommendation models. This comprehensive integration enhances understanding of citation context, leading to the generation of meaningful and actionable recommendations with precise citation alignment. The system also focuses on being explainable and transparent, providing users with explanations and visualizations for recommendations. This approach promotes user understanding and confidence, offering informative and persuasive suggestions while enhancing research reproducibility. ICA-CRMAS has proven its effectiveness through real-world evaluations and positive user feedback, addressing information overload and advancing research discovery. Future improvements include integrating powerful transformers for even more personalized recommendations and increased accuracy. The system aims to expand its coverage by incorporating additional high-quality digital libraries. Furthermore, an adaptive learning mechanism is in development for continual improvement. In conclusion, ICA-CRMAS surpasses the typical role of a recommender system; it acts as an intelligent research companion, guiding researchers through the evolving landscape of scientific knowledge with precision, relevance, context-awareness, and a commitment to continuous improvement. As we push its boundaries, ICA-CRMAS promises to remain a steadfast guide, empowering researchers to navigate the information ocean with confidence and uncover hidden gems that drive groundbreaking discoveries.

References

[1] 2020. Microsoft Academic. https://academic.microsoft.com/home. Accessed 13 August 2020.
[2] n.d.. Base-Search. Retrieved June 4, 2024, from https://www.base-search.net/.

[3] n.d.. Google Scholar. Retrieved June 4, 2024, from https://scholar.google.com/citations.
[4] Zafar Ali, Pavlos Kefalas, Khan Muhammad, Bahadar Ali, and Muhammad Imran. 2020. Deep learning in citation recommendation models survey. *Expert Systems with Applications* 162 (2020), 113790.
[5] Chandra Bhagavatula, Sergey Feldman, Russell Power, and Waleed Ammar. 2018. Content-based citation recommendation. *arXiv preprint arXiv:1802.08301* (2018).
[6] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
[7] Betül Bulut, Esra Gündoğan, Buket Kaya, Reda Alhajj, and Mehmet Kaya. 2020. User's research interests based paper recommendation system: a deep learning approach. *Putting social media and networking data in practice for education, planning, prediction and recommendation* (2020), 117–130.
[8] Xiaoyan Cai, Junwei Han, Wenjie Li, Renxian Zhang, Shirui Pan, and Libin Yang. 2018. A three-layered mutually reinforced model for personalized citation recommendation. *IEEE transactions on neural networks and learning systems* 29, 12 (2018), 6026–6037.
[9] Xiaoyan Cai, Junwei Han, and Libin Yang. 2018. Generative adversarial network based heterogeneous bibliographic network representation for personalized citation recommendation. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 32.
[10] Xiaoyan Cai, Nanxin Wang, Libin Yang, and Xin Mei. 2022. Global-local neighborhood based network representation for citation recommendation. *Applied Intelligence* (2022), 1–18.
[11] Xiaoyan Cai, Yu Zheng, Libin Yang, Tao Dai, and Lantian Guo. 2018. Bibliographic network representation based personalized citation recommendation. *IEEE Access* 7 (2018), 457–467.
[12] Jie Chen, Yang Liu, Shu Zhao, and Yanping Zhang. 2019. Citation recommendation based on weighted heterogeneous information network containing semantic linking. In *2019 IEEE international conference on multimedia and expo (ICME)*. IEEE, 31–36.
[13] Xi Chen, Huan-jing Zhao, Shu Zhao, Jie Chen, and Yan-ping Zhang. 2019. Citation recommendation based on citation tendency. *Scientometrics* 121, 2 (2019), 937–956.
[14] Xi Chen, Huan-jing Zhao, Shu Zhao, Jie Chen, and Yan-ping Zhang. 2019. Citation recommendation based on citation tendency. *Scientometrics* 121, 2 (2019), 937–956.
[15] Gohar Rehman Chughtai, Jia Lee, Asif Kabir, Rashid Abbasi, and Muhammad Arshad Shehzad Hassan. 2018. Ontology-based academic article recommendation. In *2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*. IEEE, 93–96.
[16] Fang Da, Gang Kou, and Yi Peng. 2022. Deep learning based dual encoder retrieval model for citation recommendation. *Technological Forecasting and Social Change* 177 (2022), 121545.
[17] Tao Dai, Jie Zhao, Dehong Li, Shun Tian, Xiangmo Zhao, and Shirui Pan. 2023. Heterogeneous deep graph convolutional network with citation relational BERT for COVID-19 inline citation recommendation. *Expert Systems with Applications* 213 (2023), 118841.
[18] Tao Dai, Li Zhu, Yaxiong Wang, and Kathleen M Carley. 2019. Attentive stacked denoising autoencoder with bi-lstm for personalized context-aware citation recommendation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 28 (2019), 553–568.
[19] Thi N Dinh, Phu Pham, Giang L Nguyen, and Bay Vo. 2023. Enhanced context-aware citation recommendation with auxiliary textual information based on an auto-encoding mechanism. *Applied Intelligence* 53, 14 (2023), 17381–17390.
[20] Zhengxiao Du, Jie Tang, and Yuhui Ding. 2019. Polar: Attention-based cnn for one-shot personalized article recommendation. In *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2018, Dublin, Ireland, September 10–14, 2018, Proceedings, Part II 18*. Springer, 675–690.
[21] Travis Ebesu and Yi Fang. 2017. Neural citation network for context-aware citation recommendation. In *Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval*. 1093–1096.
[22] Travis Ebesu and Yi Fang. 2017. Neural citation network for context-aware citation recommendation. In *Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval*. 1093–1096.
[23] Michael Färber and Adam Jatowt. 2020. Citation recommendation: approaches and datasets. *International Journal on Digital Libraries* 21, 4 (2020), 375–405.
[24] Michael Färber and Ashwath Sampath. 2020. Hybridcite: A hybrid model for context-aware citation recommendation. In *Proceedings of the ACM/IEEE joint conference on digital libraries in 2020*. 117–126.
[25] Michael Färber, Alexander Thiemann, and Adam Jatowt. 2018. To cite, or not to cite? Detecting citation contexts in text. In *Advances in Information Retrieval: 40th European Conference on IR Research, ECIR 2018, Grenoble, France, March 26-29, 2018, Proceedings 40*. Springer, 598–603.
[26] Soumyajit Ganguly and Vikram Pudi. 2017. Paper2vec: Combining graph and text information for scientific paper representation. In *Advances in Information Retrieval: 39th European Conference on IR Research, ECIR 2017, Aberdeen, UK, April 8-13, 2017, Proceedings 39*. Springer, 383–395.
[27] Lantian Guo, Xiaoyan Cai, Haohua Qin, Fei Hao, and Sensen Guo. 2022. A content-sensitive citation representation approach for citation recommendation. *Journal of Ambient Intelligence and Humanized Computing* (2022), 1–12.
[28] Michael Gusenbauer. 2019. Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics* 118, 1 (2019), 177–214.

[29] Donglin Hu, Huifang Ma, Yuhang Liu, and Xiangchun He. 2020. Scientific paper recommendation using author's dual role citation relationship. In *Intelligent Information Processing X: 11th IFIP TC 12 International Conference, IIP 2020, Hangzhou, China, July 3–6, 2020, Proceedings 11*. Springer, 121–132.

[30] Shi Hui, Ma Wei, Zhang XiaoLiang, Jiang JunYan, Liu YanBing, and Chen ShuJuan. 2020. A hybrid paper recommendation method by using heterogeneous graph and metadata. In *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 1–8.

[31] Chanwoo Jeong, Sion Jang, Eunjeong Park, and Sungchul Choi. 2020. A context-aware citation recommendation model with BERT and graph convolutional networks. *Scientometrics* 124 (2020), 1907–1922.

[32] Zhuoren Jiang, Yue Yin, Liangcai Gao, Yao Lu, and Xiaozhong Liu. 2018. Cross-language citation recommendation via hierarchical representation learning on heterogeneous graph. In *The 41st international ACM SIGIR conference on research & development in information retrieval*. 635–644.

[33] Anita Khadka and Petr Knoth. 2018. Using citation-context to reduce topic drifting on pure citation-based recommendation. In *Proceedings of the 12th ACM conference on recommender systems*. 362–366.

[34] Binh Thanh Kieu, Inigo Jauregi Unanue, Son Bao Pham, Hieu Xuan Phan, and Massimo Piccardi. 2021. Learning neural textual representations for citation recommendation. In *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE, 4145–4152.

[35] Xiangjie Kong, Mengyi Mao, Wei Wang, Jiaying Liu, and Bo Xu. 2018. VOPRec: Vector representation learning of papers with text information and structural identity for recommendation. *IEEE Transactions on Emerging Topics in Computing* 9, 1 (2018), 226–237.

[36] Xinyi Li, Yifan Chen, Benjamin Pettit, and Maarten De Rijke. 2019. Personalised reranking of paper recommendations using paper content and user behavior. *ACM Transactions on Information Systems (TOIS)* 37, 3 (2019), 1–23.

[37] Haifeng Liu, Xiangjie Kong, Xiaomei Bai, Wei Wang, Teshome Megersa Bekele, and Feng Xia. 2015. Context-based collaborative filtering for citation recommendation. *Ieee Access* 3 (2015), 1695–1703.

[38] Yonghe Lu, Meilu Yuan, Jiaxin Liu, and Minghong Chen. 2023. Research on semantic representation and citation recommendation of scientific papers with multiple semantics fusion. *Scientometrics* 128, 2 (2023), 1367–1393.

[39] Xiao Ma and Ranran Wang. 2019. Personalized scientific paper recommendation based on heterogeneous graph representation. *IEEE Access* 7 (2019), 79887–79894.

[40] Xiao Ma and Ranran Wang. 2019. Personalized scientific paper recommendation based on heterogeneous graph representation. *IEEE Access* 7 (2019), 79887–79894.

[41] Xiao Ma, Yin Zhang, and Jiangfeng Zeng. 2019. Newly published scientific papers recommendation in heterogeneous information networks. *Mobile Networks and Applications* 24 (2019), 69–79.

[42] Lu Meilian, Wei Xudan, Gao Jie, and Shi Yan. 2015. Ahits-upt: A high quality academic resources recommendation method. In *2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity)*. IEEE, 507–512.

[43] Phu Pham, Hieu Le, Nguyen Thanh Tam, and Quang-Dieu Tran. 2023. A Graph-Based Topic Modeling Approach to Detection of Irrelevant Citations. *Vietnam Journal of Computer Science* 10, 02 (2023), 197–216.

[44] Chanathip Pornprasit, Xin Liu, Pattararat Kiattipadungkul, Natthawut Kertkeidkachorn, Kyoung-Sook Kim, Thanapon Noraset, Saeed-Ul Hassan, and Suppawong Tuarob. 2022. Enhancing citation recommendation using citation network embedding. *Scientometrics* (2022), 1–32.

[45] Nazmus Sakib, Rodina Binti Ahmad, and Khalid Haruna. 2020. A collaborative approach toward scientific paper recommendation using citation context. *IEEE Access* 8 (2020), 51246–51255.

[46] Kazunari Sugiyama and Min-Yen Kan. 2015. A comprehensive evaluation of scholarly paper recommendation using potential citation papers. *International Journal on Digital Libraries* 16, 2 (2015), 91–109.

[47] Yifan Wang, Yiping Song, Shuai Li, Chaoran Cheng, Wei Ju, Ming Zhang, and Sheng Wang. 2022. Disencite: Graph-based disentangled representation learning for context-specific citation generation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36. 11449–11458.

[48] Qianqian Xie, Yutao Zhu, Jimin Huang, Pan Du, and Jian-Yun Nie. 2021. Graph neural collaborative topic model for citation recommendation. *ACM Transactions on Information Systems (TOIS)* 40, 3 (2021), 1–30.

[49] Cheng Yang, Zhiyuan Liu, Deli Zhao, Maosong Sun, and Edward Y Chang. 2015. Network representation learning with rich text information.. In *IJCAI*, Vol. 2015. 2111–2117.

[50] Libin Yang, Zeqing Zhang, Xiaoyan Cai, and Tao Dai. 2019. Attention-Based Personalized Encoder-Decoder Model for Local Citation Recommendation. *Computational Intelligence and Neuroscience* 2019, 1 (2019), 1232581.

[51] Libin Yang, Zeqing Zhang, Xiaoyan Cai, and Lantian Guo. 2019. Citation recommendation as edge prediction in heterogeneous bibliographic network: a network representation approach. *IEEE Access* 7 (2019), 23232–23239.

[52] Libin Yang, Yu Zheng, Xiaoyan Cai, Hang Dai, Dejun Mu, Lantian Guo, and Tao Dai. 2018. A LSTM based model for personalized context-aware citation recommendation. *IEEE access* 6 (2018), 59618–59627.

[53] Yang Zhang and Qiang Ma. 2020. Citation recommendations considering content and structural context embedding. In *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)*. IEEE, 1–7.

[54] Ye Zhang, Libin Yang, Xiaoyan Cai, and Hang Dai. 2018. A novel personalized citation recommendation approach based on GAN. In *Foundations of Intelligent Systems: 24th International Symposium, ISMIS 2018, Limassol, Cyprus, October 29–31, 2018, Proceedings 24*. Springer, 268–278.
[55] Muhammad Azeem Abbas, Saheed Ajayi, Muhammad Bilal, Ade Oyegoke, Maruf Pasha, and Hafiz Tauqeer Ali. 2022. A deep learning approach for context-aware citation recommendation using rhetorical zone classification and similarity to overcome cold-start problem. *Journal of Ambient Intelligence and Humanized Computing* (2022), 1–15.
[56] Trapit Bansal, David Belanger, and Andrew McCallum. 2016. Ask the gru: Multi-task learning for deep text recommendations. In *proceedings of the 10th ACM Conference on Recommender Systems*. 107–114.
[57] Lukas Galke, Florian Mai, Iacopo Vagliano, and Ansgar Scherp. 2018. Multi-modal adversarial autoencoders for recommendations of citations and subject labels. In *Proceedings of the 26th conference on user modeling, adaptation and personalization*. 197–205.
[58] Shashank Gupta and Vasudeva Varma. 2017. Scientific article recommendation by using distributed representations of text and graph. In *Proceedings of the 26th international conference on world wide web companion*. 1267–1268.
[59] Damien Hanyurwimfura, Liao Bo, Vincent Havyarimana, Dennis Njagi, and Faustin Kagorora. 2015. An effective academic research papers recommendation for non-profiled users. *International Journal of Hybrid Information Technology* 8, 3 (2015), 255–272.
[60] Hebatallah A Mohamed Hassan. 2017. Personalized research paper recommendation using deep learning. In *Proceedings of the 25th conference on user modeling, adaptation and personalization*. 327–330.
[61] Wenyi Huang, Zhaohui Wu, Chen Liang, Prasenjit Mitra, and C Giles. 2015. A neural probabilistic model for context based citation recommendation. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 29.
[62] Chanwoo Jeong, Sion Jang, Eunjeong Park, and Sungchul Choi. 2020. A context-aware citation recommendation model with BERT and graph convolutional networks. *Scientometrics* 124 (2020), 1907–1922.
[63] Yuta Kobayashi, Masashi Shimbo, and Yuji Matsumoto. 2018. Citation recommendation using distributed representation of discourse facets in scientific articles. In *Proceedings of the 18th ACM/IEEE on joint conference on digital libraries*. 243–251.
[64] Joonseok Lee, Kisung Lee, Jennifer G Kim, and Sookyung Kim. 2015. Personalized academic paper recommendation system. *SRS'15* (2015).
[65] Patricia Maraver, Rubén Armañanzas, Todd A Gillette, and Giorgio A Ascoli. 2019. PaperBot: open-source web-based search and metadata organization of scientific literature. *BMC bioinformatics* 20 (2019), 1–13.
[66] Dejun Mu, Lantian Guo, Xiaoyan Cai, and Fei Hao. 2017. Query-focused personalized citation recommendation with mutually reinforced ranking. *IEEE Access* 6 (2017), 3107–3119.
[67] Fernando Ortega, Jesús Bobadilla, Abraham Gutiérrez, Remigio Hurtado, and Xin Li. 2018. Artificial intelligence scientific documentation dataset for recommender systems. *IEEE access* 6 (2018), 48543–48555.
[68] Malte Ostendorff, Nils Rethmeier, Isabelle Augenstein, Bela Gipp, and Georg Rehm. 2022. Neighborhood contrastive learning for scientific document representations with citation embeddings. *arXiv preprint arXiv:2202.06671* (2022).
[69] Ritu Sharma, Dinesh Gopalani, and Yogesh Meena. 2017. Concept-based approach for research paper recommendation. In *International Conference on Pattern Recognition and Machine Intelligence*. Springer, 687–692.
[70] Jie Tang and Jing Zhang. 2009. A discriminative approach to topic-based citation recommendation. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 572–579.
[71] Han Tian and Hankz Hankui Zhuo. 2017. Paper2vec: Citation-context based document distributed representation for scholar recommendation. *arXiv preprint arXiv:1703.06587* (2017).
[72] Wei Wang, Tao Tang, Feng Xia, Zhiguo Gong, Zhikui Chen, and Huan Liu. 2020. Collaborative filtering with network representation learning for citation recommendation. *IEEE Transactions on Big Data* 8, 5 (2020), 1233–1246.
[73] Kyle Williams, Jian Wu, and C Lee Giles. 2014. Simseerx: a similar document search engine. In *Proceedings of the 2014 ACM symposium on Document engineering*. 143–146.
[74] Jun Yin and Xiaoming Li. 2017. Personalized citation recommendation via convolutional neural networks. In *Web and Big Data: First International Joint Conference, APWeb-WAIM 2017, Beijing, China, July 7–9, 2017, Proceedings, Part II 1*. Springer, 285–293.

A Related works appendix

Table 13 provides a comprehensive comparison between various citation recommendation models. The table includes information on filtering methods, data models, and problems faced by each model, highlighting the differences and strengths of each approach.

Model		Filtering method								Information used								Data model			Problem faced		
		Collaborative	Content-based	Graph-based	Machine Learning	Deep learning	Multimodal learning	Hybrid	Context-aware	Paper contents	Tags/Keywords	Author profile	Venue information	Citation network	Social network	Rating	Images	Matrix-based	Graph-based	Ontology-based	Sparsity	Cold-start	Accuracy
10	Kieu(2020) [34]	✓	-	-	-	✓	-	-	-	✓	-	-	✓	✓	-	-	-	-	✓	✓	-	-	✓
13	RBM-CS ([70])	-	-	-	-	✓	-	-	-	-	-	✓	-	✓	-	-	-	-	✓	-	-	-	-
1	Simseerx ([73])	-	✓	-	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-
2	Lee(2015)[64]	✓	-	-	-	-	-	-	-	✓	-	-	-	-	✓	-	-	✓	-	-	-	-	-
3	LIU(2015) [37]	✓	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	✓	-	-	-	-	-
4	Njagi(2015) [59]	✓	-	-	-	-	-	-	-	✓	✓	-	-	-	-	-	-	-	-	-	-	✓	✓
33	NPM ([61])	-	✓	-	✓		-	-	✓	✓	✓	✓	-	-	-	-	-	✓	-	-	-	-	-
5	Sugiyama(2015) [46]	✓	-	-	✓	-	-	-	-	✓	-	-	-	✓	-	-	-	✓	-	-	✓	-	-
22	ML-DTR ([56])	✓	-	-	-	✓	-	-	-	✓	✓	✓	-	-	-	✓	-	✓	-	-	✓	✓	✓
23	SAR ([58])	✓	-	✓	-	-	-	-	-	-	-	-	✓	-	✓	-	-	-	✓	-	-	-	-
24	BERT-GCN ([62])	-	-	✓	-	✓	-	-	✓	✓	-	✓	-	✓	-	-	-	-	✓	-	-	-	✓
25	NCN ([22])	-	✓	-	-	✓	-	-	✓	✓	-	-	-	✓	✓	-	-	-	-	-	-	-	-
26	PPR-DL ([60])	-	-	-	-	✓	-	-	-	✓	-	✓	-	-	-	-	-	✓	-	-	-	-	✓
34	Paper2vec ([71])	-	-	-	✓	-	-	-	-	✓	-	-	-	✓	-	-	-	✓	-	-	-	-	-
36	p-CNN ([74])	-	-	-	-	✓	-	-	-	✓	-	✓	-	-	✓	-	-	✓	-	-	-	-	✓
39	CPR ([69])	-	-	-	-	-	-	-	-	✓	-	✓	-	-	-	-	-	✓	-	-	-	-	✓
6	ORTEGA(2017) [67]	✓	-	-	-	-	-	-	-	✓	✓	-	-	✓	-	✓	-	-	-	-	✓	-	✓
7	PaperBot ([65])	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	✓	-	-	-	-	-
14	MMRQ ([66])	-	-	-	-	-	-	-	-	✓	✓	✓	-	✓	-	-	-	-	✓	-	-	-	✓
15	MAAE ([57])	✓	-	-	-	-	-	-	-	✓	✓	✓	-	✓	-	✓	-	✓	-	-	✓	-	-
20	RDF-CR ([63])	-	✓	-	-	-	-	-	-	-	-	✓	-	✓	-	-	-	✓	✓	-	-	-	✓

Table 13. Comparison between citation recommendation models.

Datasets	Papers	Authors	Venues	Citation relations	Citation context	Tags	Key phrases	Title	Abstract	Full text	Year info	Ratings	Quality Metrics	Images	Release
OpenCorpus	60,90,000	80,30,000	23,672	✓	-	-	✓	✓	✓	-	✓	-	-	-	2018
RARD II	24,00000	-	-	-	-	-	-	-	-	-	-	✓	-	-	2018
DBLP	4,107,340	318,406	23,709	✓	-	-	-	✓	✓	-	✓	-	-	-	2019
ACL Anthology	23766	18862	373	✓	-	-	-	✓	✓	✓	-	-	-	-	2016
RefSeer	855,735	-	-	✓	✓	-	✓	-	✓	✓	✓	-	-	-	2015
Aminer	3,680,007	212,567	12,770	✓	-	-	-	✓	✓	-	✓	-	-	-	2017
arXiv CS	90,278	269,194	1,489	✓	✓	-	✓	✓	✓	-	✓	-	-	-	2018
Scholarly Dataset	100351	50	-	✓	-	-	-	-	✓	✓	-	-	-	-	2013
Onto-CCR	**100,000**	**320,154**	**7068**	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	**2023**

Table 14. Datasets specifications.

B Datasets specifications

Table 14 outlines the specifications of various datasets used in our study. This table includes detailed information on the number of papers, authors, venues, and additional attributes such as citation relations, citation context, and key phrases for each dataset.

C Illustrative Examples

The usage of ICA-CRMAS starts with user authentication via a login interface, granting access to personalized profiles for tailored recommendations. Users can customize their searches via a dedicated configuration panel, specifying search sentences, the desired number of papers for recommendation, and publication date ranges to filter articles. Figure 12 provides a visual representation of this configuration panel. Following search criteria definition, ICA-CRMAS initiates web crawling to locate relevant scholarly articles. Once retrieved, the system activates the recommendation process, leveraging its capabilities to analyze and recommend the most relevant papers for the user's research needs. The recommended papers are then presented in a user-friendly interface (Figure 13), displaying essential details for each paper to facilitate rapid assessment of their relevance. Furthermore, ICA-CRMAS constructs a personalized database of relevant papers over time, accumulating recommendations and enabling easy access to previously identified citations for ongoing research activities.

Our system leverages an ontology to represent knowledge about research articles, authors, topics, and citation connections. Embedded within this ontology are Semantic Web Rule Language (SWRL) rules that govern the reasoning process. These rules provide the foundation for generating insightful explanations for recommended citations, crafting detailed justifications for each paper. The figure (14) presents an example of our system's explanation for the first recommended paper. This explanation feature offers several functionalities that further enhance the transparency of the recommendation process:

- The system displays the Article Node Power for each article, arranged in descending order from highest to lowest. Additionally, it provides insights into factors influencing the selection of the most relevant papers, such as PW, CBP, PMA, and others.
- The system offers a visualization tool for displaying recommended articles in a graph format. This functionality enables users to interact with the ontology graph, facilitating exploration of various ontology sections, zooming in and out, and inspecting node and edge details. This feature aids users in tracing specific papers, understanding the significant relationships and parameters contributing to their recommendation, and identifying the navigational path followed by the system to recommend each paper.

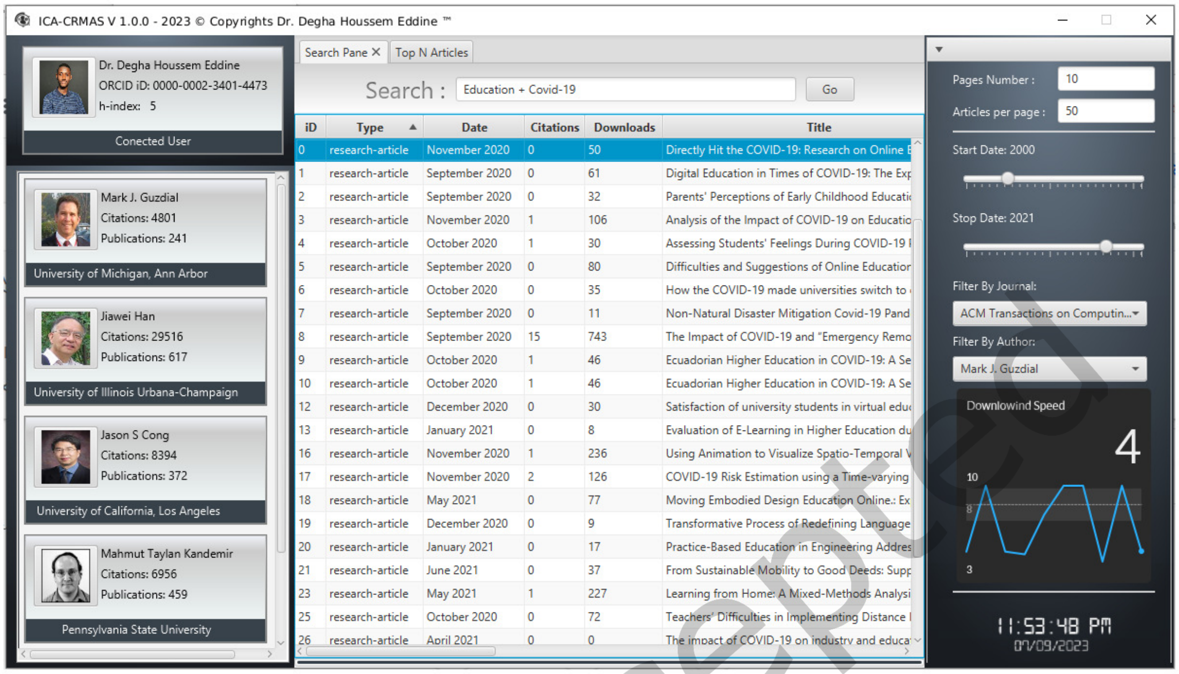


Fig. 12. illustrates the top ten article recommendations generated by ICA-CRMAS.

- Our system leverages an ontology and Semantic Web Rule Language (SWRL) rules to generate comprehensive explanations for recommendations. It presents all the rules underlying the rationale for recommending each paper. Below are examples of the advanced SWRL rules utilized:

```
Number_Of_Shared_Keywords(?p1, ?p2, ?count) ^ (count > (
    ↪ Threshold_Keyword_Similarity * Total_Keywords(?p1))) -> contributesTo(
    ↪ Explanation(?p2, "Significantly shares keywords with P1 (weight: ?
    ↪ weight)."), ?weight).
Weight = Number_Of_Shared_Keywords(?p1, ?p2, ?) / Total_Keywords(?p1)
```

```
TopicSimilarity(?p1, ?p2, ?similarityScore) ^ (similarityScore >
    ↪ Threshold_Topic_Alignment) -> Explanation(?p2, "Highly aligned with P1
    ↪  in terms of topic (weight: ?weight)."), ?weight).
```

```
Author(?p2, ?author) ^ ExpertInField(?author, ?field) ^ Topic(?p1, ?topic) ^
    ↪ SubTopicOf(?topic, ?field) -> Explanation(?p2, "Authored by an expert
    ↪ in a relevant subfield (weight: 0.7).").
```

```
PublicationYear(?p1, ?year1) ^ PublicationYear(?p2, ?year2) ^ abs(diff(?year1
    ↪ , ?year2)) <= Threshold_Publication_Year_Difference -> contributesTo(
    ↪ Explanation(?p2, "Recently published in the same field (weight: 0.3)
    ↪ ."))
```

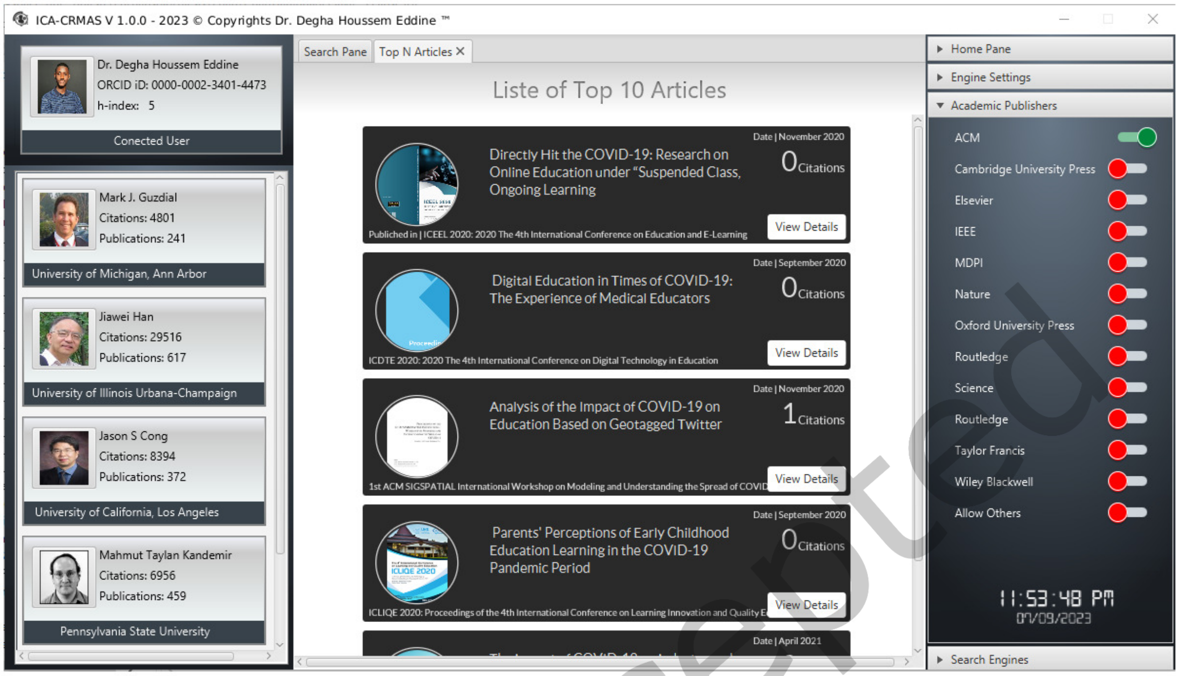


Fig. 13. showcases the configuration panel and search results presentation within ICA-CRMAS.

D Description of Key Classes in Onto-CCR

The modeling of concepts, attributes, and relationships within Onto-CCR is conducted using the Web Ontology Language (OWL). These concepts and relationships are defined through hyperlinks, guiding the OWL description of each concept. The following sections provide introductions to key Onto-CCR concepts. Since the ontology concepts are expressed in the Description Logic language (DL). We additionally define the notations employed in this section:

- The expressions "Name1.Name2" and "Name1 ⊑ Name2" **indicate that "Name1" is a subclass of "Name2".**
- The term "Onto-CCR" **refers to** the name of our ontology, Onto-CCR (smart building ontology).
- The symbol "∩" **represents** the logical operator "And".
- The expression "= Nb Name(xsd:type)" **defines** the data-type properties of a class. The number "Nb" specifies the minimum number of data-type properties, "Name" refers to the property name, and "(xsd:type)" denotes the property type.
- The complete expression "∃ R: name" **indicates that all individuals must have at least one relationship with the concept "name" through the object-property "R".**
- A period "." **marks the end of the expression.**

D.0.1 Paper class. The "Paper" class, serving as a cornerstone within our ontology, encapsulates the representation of scientific papers. Operating as a distinctive form of academic discourse, a research paper intricately examines a specific subject matter, subjecting it to critical evaluation and interpretation grounded in empirical evidence. The "Paper" class encompasses a spectrum of attributes, including title, abstract, keywords, and temporal

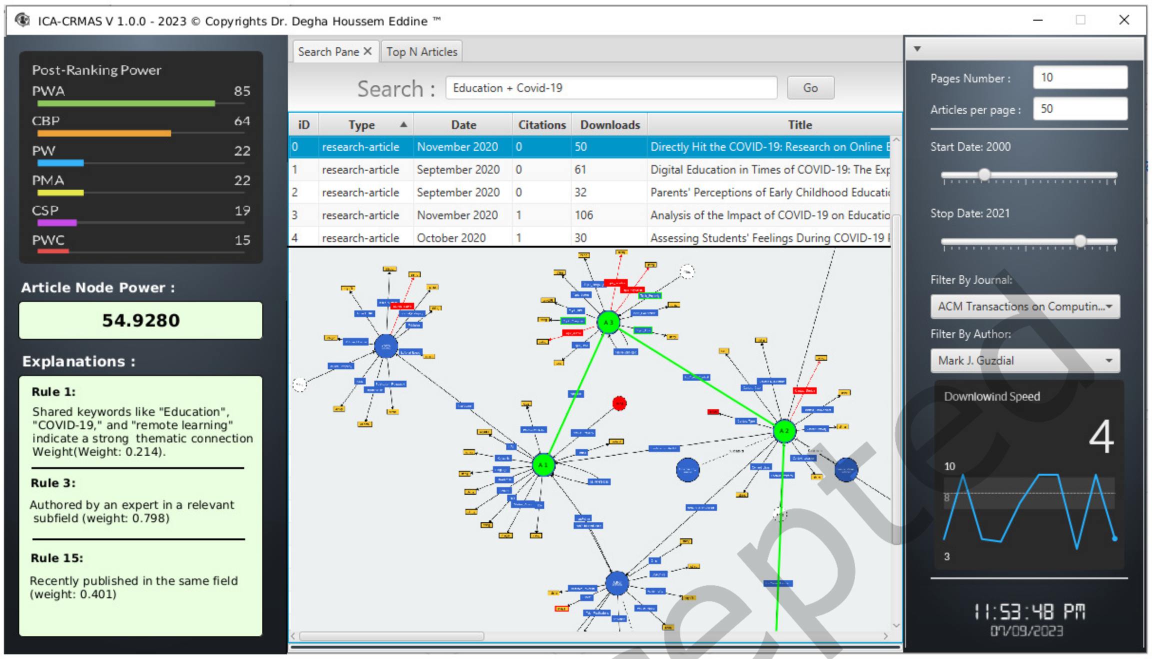


Fig. 14. An illustrative example demonstrating the explanation feature for the recommended paper using SWRL rules

dimensions such as received date, accepted date, and publication date.

Moreover, each paper establishes connections with essential elements, including authors, its chronological trajectory, references, topics, citation metrics, and other pertinent details. This interconnected web of information collaboratively constructs a comprehensive and nuanced portrayal of the paper. The ensuing comprehensive depiction of the "Paper"(4) class within our ontology is expounded upon in further detail below.

$$\begin{aligned}
&\textbf{Paper} \sqsubseteq Onto-CCR.Owl \quad \cap \\
&= 1 \quad \textbf{Article Identifier}(xsd:String) \quad \cap \\
&= 1 \quad \textbf{Content Text Data}(xsd:String)[Title, Abstract, Keywords, Full_Text, Highlights] \quad \cap \\
&= 1 \quad \textbf{Received Date}(xsd:Date) \quad \cap \\
&= 1 \quad \textbf{Revised Date}(xsd:Date) \quad \cap \\
&= 1 \quad \textbf{Accepted Date}(xsd:Date) \quad \cap \\
&= 1 \quad \textbf{Available online Date}(xsd:Date) \quad \cap \\
&= 1 \quad \textbf{Publication Date}(xsd:Date) \quad \cap \\
&= Multi \quad \textbf{Uniform Resource Indicator}(xsd:String)[DoiLink, LinkofArticle, self-uri] \quad \cap \\
&= Multi \quad \textbf{Supplement Information}(xsd:Object) \quad \cap \\
&= 1 \quad \textbf{Issue}(xsd:String)[IssueIdentifier, IssueTitle, IssueNumber] \quad \cap \\
&= 1 \quad \textbf{Volume}(xsd:Integer)[VolumeIdentifier, VolumeNumber] \quad \cap \\
&= Multi \quad \textbf{Copyright}(xsd:String)[CopyrightStatement, CopyrightYear, LicenseInformation, Permissions] \quad \cap \\
&= Multi \quad \textbf{Tags}(xsd:String) \quad \cap \\
&\exists \quad \textbf{HasAuthors} : Onto-CCR.Author \quad \cap \\
&\exists \quad \textbf{HasCorresponding Author} : Onto-CCR.Author \quad \cap \\
&\exists \quad \textbf{PublishedIn} : Onto-CCR.Journal \quad \cap \\
&\exists \quad \textbf{HasContext} : Onto-CCR.Context \quad \cap \\
&\exists \quad \textbf{HasHistory} : Onto-CCR.History \quad \cap \\
&\exists \quad \textbf{HasMaterials} : Onto-CCR.SupplementaryMaterial \quad \cap \\
&\exists \quad \textbf{HasTopics} : Onto-CCR.Topic \quad \cap \\
&\exists \quad \textbf{HasReferences} : Onto-CCR.Paper \quad \cap \\
&\exists \quad \textbf{CitedBy} : Onto-CCR.Paper \quad \cap \\
&\exists \quad \textbf{HasQualityMetrics} : Onto-CCR.Qualitymetrics[Citation] \quad \cap
\end{aligned} \tag{4}$$

D.0.2 Author class. The "Author" class in our ontology represents the individuals behind a paper's creation. It delves beyond mere identification, capturing a comprehensive author profile through details like name, affiliation, and even a brief bio. This richness extends to their contributions, impact metrics, and other relevant information, providing a nuanced understanding of their role in the scholarly landscape. The complete definition of the "Author"(5) class within our ontology is presented below:

$$\begin{aligned}
&\textbf{Author} \sqsubseteq \text{Onto-CCR.Human} \cap \\
&\quad = 1\textbf{Identifiers}[\text{orcid}](\text{xsd:String}) \cap \\
&\quad = 1\textbf{Name}(\text{xsd:String}) \cap \\
&\quad = 1\textbf{LastName}(\text{xsd:String}) \cap \\
&\quad = 1\textbf{Birthday}(\text{xsd:Date}) \cap \\
&\quad = 1\textbf{Age}(\text{xsd:int}) \cap \\
&\quad = 1\textbf{Gender}(\text{xsd:String}) \cap \\
&\quad = 1\textbf{Contact Information}(\text{xsd:String})[\text{Email, Phone Number}] \cap \\
&\quad = 1\textbf{Short Biography}(\text{xsd:String}) \cap \\
&\exists\textbf{HasProfile} : \text{Onto-CCR.Profile.Author} \cap \\
&\exists\textbf{HasAffiliation} : \text{Onto-CCR.Profile.Affiliation} \cap \\
&\exists\textbf{Lives in Country} : \text{Onto-CCR.Country} \cap \\
&\exists\textbf{HasSocial Media Profile} : \text{Onto-CCR.Social Media} \cap \\
&\exists\textbf{HasMetrics} : \text{Onto-CCR.Metrics}[\text{H-Index, Publications Count, Citations, Reviews}] \cap \\
&\exists\textbf{HasPublications} : \text{Onto-CCR.Paper} \cap \\
&\exists\textbf{HasContributions} : \text{Onto-CCR.Contributions} \cap \\
&\exists\textbf{HasCo-Authors} : \text{Onto-CCR.Author} \cap \\
&\exists\textbf{HasFollowers} : \text{Onto-CCR.Author} \cap \\
&\exists\textbf{Following} : \text{Onto-CCR.Author} \cap
\end{aligned} \tag{5}$$

D.0.3 Author Profile. Our Onto-CCR ontology features a comprehensive 'Author Profile' class, dedicated to capturing the multifaceted nature of author profiles. Each author within the ontology is enriched with diverse attributes encompassing interests, preferences, skills, roles, activities, and other relevant properties. This detailed information contributes to a nuanced understanding of individual user profiles, providing valuable insights beyond mere identification. The comprehensive delineation of the "Author Profile"(6) class within our ontology is provided below:

$$\begin{aligned}
&\textbf{Author Profile} \sqsubseteq \text{Onto-CCR.Owl.Profile}\cap \\
&\quad = 1\textbf{ProfileID}(\text{xsd:String})\cap \\
&\quad = 1\textbf{Popularity}(\text{xsd:Float})\cap \\
&\quad \exists\textbf{HasRole} : \text{Onto-CCR.Role}\cap \\
&\quad \exists\textbf{HasSkills} : \text{Onto-CCR.Skills}\cap \\
&\quad \exists\textbf{HasState} : \text{Onto-CCR.HumanState}\cap \\
&\quad \exists\textbf{HasActivity} : \text{Onto-CCR.Activity}\cap \\
&\quad \exists\textbf{HasBehavior} : \text{Onto-CCR.Behavior}\cap \\
&\quad \exists\textbf{HasProperty} : \text{Onto-CCR.Property}\cap \\
&\quad \exists\textbf{HasIdentity} : \text{Onto-CCR.Identity}\cap \\
&\quad \exists\textbf{HasInterest} : \text{Onto-CCR.Interest}\cap \\
&\quad \exists\textbf{HasCalendar} : \text{Onto-CCR.Calendar}\cap \\
&\quad \exists\textbf{HasPreferences} : \text{Onto-CCR.Preferences}
\end{aligned} \tag{6}$$

D.0.4 Journal. The "Journal" class in our Onto-CCR ontology represents scholarly publications where authors disseminate their research findings. It encompasses various data types describing each journal, including identifiers (ISSN, EISSN), title, evaluation metrics, volumes, publication time, and homepage. Additionally, the "Journal" class establishes connections with other classes within the ontology. For example, each journal has a defined scope, publishes articles, has an editorial board responsible for overseeing the publication process, and is managed by a publisher. The detailed exposition of the "Journal"(7) class in our ontology is articulated below:

$$\begin{aligned}
&\textbf{Journal} \sqsubseteq \text{Onto-CCR.Owl}\cap \\
&\quad = 1\textbf{Identifiers}[\text{ISSN, EISSN}](\text{xsd:String})\cap \\
&\quad = 1\textbf{Title}(\text{xsd:String})\cap \\
&\quad = \text{Multi}\textbf{Metrics}[\text{CiteScore, Impact Factor, Citable Items, H5-Index}](\text{xsd:String})\cap \\
&\quad = \text{Multi}\textbf{Volume}(\text{xsd:String})\cap \\
&\quad = 1\textbf{Founded}(\text{xsd:String})\cap \\
&\quad = 1\textbf{Publication Time}(\text{xsd:Time})\cap \\
&\quad = 1\textbf{Homepage}(\text{xsd:String})\cap \\
&\exists\textbf{HasScope} : \text{Onto-CCR.Scope}\cap \\
&\exists\textbf{HasEditorialBoard} : \text{Onto-CCR.EditorialBoard}\cap \\
&\exists\textbf{HasAbstractingIndexing} : \text{Onto-CCR.AbstractingIndexing}\cap \\
&\exists\textbf{HasPapers} : \text{Onto-CCR.Paper}\cap \\
&\exists\textbf{HasPublisher} : \text{Onto-CCR.Publisher}\cap \\
&\exists\textbf{HasTopics} : \text{Onto-CCR.Topics}\cap
\end{aligned} \tag{7}$$

D.0.5 Metrics. Our Onto-CCR ontology transcends limitations of conventional research quality evaluation by offering a rich tapestry of metrics encompassing diverse entities beyond just journals and papers. This comprehensive framework delves into the impact of research across various domains, including:

- Authors: Capture the true influence of researchers, not just through basic citations, but also through sophisticated indicators like Category Normalised Citation Impact (CNCI) and Field-Weighted Citation Impact (FWCI), revealing their standing within specific areas.

- Papers: Move beyond simple citation counts to assess a paper's individual merit. Onto-CCR utilizes Article Influence Score, Field-Weighted Citation Impact, and mentions across diverse platforms like blog posts and reviews to paint a holistic picture of a paper's reach and significance.
- Journals: Evaluate the standing and reputation of scholarly journals with established indicators like Journal Impact Factor and CiteScore, while also incorporating emerging metrics like Eigenfactor for a more nuanced understanding of their influence within the broader research landscape.
- Data and Social Media: Onto-CCR extends its reach beyond traditional research entities, providing metrics to assess the reliability and trustworthiness of data (accuracy, validity, uniqueness) and gauge the public engagement and reception of research findings on social media platforms (likes, shares).

By combining these diverse metrics, Onto-CCR empowers our system to make informed decisions based on a holistic understanding of research quality and impact. This comprehensive framework moves beyond traditional limitations, offering a valuable tool for evaluating research across various entities and contexts. Each category within the "Metrics"(8,9,10,11,12) class is further enriched by a diverse range of metrics, tailored to provide a comprehensive and insightful evaluation:

$$\begin{aligned}
&\textbf{Author Metrics} \sqsubseteq Onto-CCR.Metrics \;\; \cap \\
&= Multi \;\; Citationsperpublication(xsd:Inetger) \;\; \cap \\
&= 1 \;\; CNCI(CategoryNormalisedCitationImpact)(xsd:Inetger) \;\; \cap \\
&= 1 \;\; FWCI(Field-WeightedCitationImpact)(xsd:Inetger) \;\; \cap \\
&= 1 \;\; h-index(xsd:Inetger) \;\; \cap \\
&= 1 \;\; Citationscount(xsd:Inetger) \;\; \cap \\
&= Multi \;\; self-citations(xsd:Inetger) \;\; \cap \\
&= 1 \;\; i10-index(xsd:Inetger) \;\; \cap
\end{aligned} \tag{8}$$

$$\begin{aligned}
&\textbf{Paper Quality Metrics} \sqsubseteq Onto-CCR.Metrics \;\; \cap \\
&= 1 \;\; ArticleInfluenceScore(xsd:Inetger) \;\; \cap \\
&= 1 \;\; Citation(xsd:Inetger) \;\; \cap \\
&= 1 \;\; Counts(xsd:Inetger) \;\; \cap \\
&= 1 \;\; Field-WeightedCitationImpact(FWCI)(xsd:Inetger) \;\; \cap \\
&= 1 \;\; CNCI(CategoryNormalisedCitationImpact)(xsd:Inetger) \;\; \cap \\
&= Multi \;\; Mentionss[blogposts, comments, reviews, Wikipediareferences, newsmedia](xsd:Inetger) \;\; \cap \\
&= Multi \;\; Captures[Mendeley, bookmarks](xsd:Inetger) \;\; \cap \\
&= Multi \;\; Usage[clicks, downloads, views, holdings](xsd:Inetger) \;\; \cap
\end{aligned} \tag{9}$$

$$\begin{aligned}
&\textbf{journal Quality Metrics} \sqsubseteq Onto-CCR.Metrics \;\; \cap \\
&= 1 \;\; JournalImpactFactor(xsd:Inetger) \;\; \cap \\
&= 1 \;\; SJR(xsd:Inetger) \;\; \cap \\
&= 1 \;\; SNIP(xsd:Inetger) \;\; \cap \\
&= 1 \;\; Citations(xsd:Inetger) \;\; \cap \\
&= 1 \;\; CiteScore(xsd:Inetger) \;\; \cap \\
&= 1 \;\; Eigenfactor(xsd:Inetger) \;\; \cap
\end{aligned} \tag{10}$$

$$\begin{aligned}
&\textbf{Data Quality Metrics} \sqsubseteq Onto-CCR.Metric \;\; \cap \\
&= 1 \;\; Accurancy(xsd:Inetger) \;\; \cap \\
&= 1 \;\; Validity(xsd:Inetger) \;\; \cap \\
&= 1 \;\; Uniquness(xsd:Inetger) \;\; \cap \\
&= 1 \;\; Timeliness(xsd:Inetger) \;\; \cap \\
&= Multi \;\; Consistency(xsd:Inetger) \;\; \cap \\
&= Multi \;\; Completness(xsd:Inetger) \;\; \cap
\end{aligned} \tag{11}$$

$$
\begin{aligned}
&\textbf{Social media Quality Metrics} \sqsubseteq Onto - CCR.Metrics \ \cap \\
&= 1 \ \ likes(xsd : Inetger) \ \cap \\
&= 1 \ \ buzzmeasure(xsd : Inetger) \ \cap \\
&= 1 \ \ attentionmeasure(xsd : Inetger) \ \cap \\
&= 1 \ \ shares(xsd : Inetger) \ \cap \\
&= Multi \ \ emojis(xsd : Inetger) \ \cap \\
&= Multi \ \ comments(xsd : Inetger) \ \cap
\end{aligned} \tag{12}
$$

D.0.6 Topic. The **Topic** class serves as the foundational framework within the Onto-CCR ontology, akin to a robust scaffold that organizes and interconnects research knowledge. It relies on two fundamental pillars:

- **Name:** A succinct label that delineates the specific domain of knowledge encapsulated by the topic.
- **Keywords:** A compilation of descriptive terms refining and broadening the scope of the topic, facilitating efficient retrieval and exploration.

However, the Topic class doesn't exist in isolation. It intricately weaves itself into a complex tapestry of relationships, establishing connections among various knowledge elements:

- **Related To:** This linkage bridges the gap between topics sharing conceptual proximity or thematic overlap, forming a network of interconnected knowledge. The rule for creating this relation is presented below:

```
Topic(?x) ^ Topic(?y) ^ differentFrom(?x, ?y) ^
HasKeywords(?paper, ?keywords) ^
StringSetContains(?keywords, ?x) ^ StringSetContains(?keywords, ?y) ->
    ↪ RelatedTo(?x, ?y)
```

Has Sub Topic and Has Super Topic: These relationships create a hierarchical structure within the knowledge domain. Visualize "Machine Learning" as a sturdy platform, with "Natural Language Processing" and "Deep Learning" branching out as specialized subfields, each building upon the foundation laid by its superclass. The rule for creating this relation is presented below:

```
Topic(?a) ^ Topic(?b) ^ differentFrom(?a, ?b) ^
Paper(?p1) ^ CitedBy(?p2, ?p1) ^ HasTopics(?p1, ?a) ^ HasTopics(?p2, ?b) ^
~HasSuperTopic(?a, ?b) -> SubTopicOf(?b, ?a)
```

```
Topic(?a) ^ Topic(?b) ^ differentFrom(?a, ?b) ^
Paper(?p1) ^ CitedBy(?p2, ?p1) ^ HasTopics(?p1, ?a) ^ HasTopics(?p2, ?b) ^
~SubTopicOf(?a, ?b) -> SuperTopicOf(?a, ?b)
```

Beyond this internal network, the Topic class extends its influence, forging vital links with other key entities in the ontology:

- **Paper:** Each research paper finds its place within the thematic framework by aligning with relevant topics, enabling efficient categorization and retrieval.
- **Author and Journal:** These entities, too, can be associated with specific topics, offering insights into individual research interests and the thematic focus of publications.

Through this intricate web of connections, the Topic class empowers Onto-CCR to navigate and analyze research knowledge with unparalleled precision. It serves as the scaffolding upholding the structure of knowledge, the map guiding exploration, and the bridge connecting diverse research elements. The comprehensive description of the "Topic"(13) class in our ontology is presented below:

$$\begin{aligned}
&Topic \sqsubseteq Onto-CCR \ \cap \\
&= 1 \ \ Name(xsd:String) \ \cap \\
&= multi \ \ Keywords(xsd:String) \ \cap \\
&\exists \ \ RelatedTo: Onto-CCR.Topic \ \cap \\
&\exists \ \ HasSubTopic: Onto-CCR.Topic \ \cap \\
&\exists \ \ HasSuperTopic: Onto-CCR.Topic \ \cap
\end{aligned} \tag{13}$$

D.0.7 Context. The **Context** class stands at the core of our Onto-CCR ontology, serving as a pivotal lens through which we comprehend and analyze scientific papers. It transcends a mere aggregation of data, evolving into a meticulously constructed framework that interweaves various dimensions of information, crafting a holistic portrayal of research work.

This intricate tapestry of context is woven from diverse threads, each contributing invaluable insights:

- **Attributes:** These encapsulate the essential characteristics of the research, including its topic, methodology, and findings.
- **Life Cycle and State:** We meticulously monitor the evolution of the context, offering a chronological timeline of its development.
- **Type:** Papers are systematically categorized based on their structure and purpose, facilitating targeted exploration.
- **Quality Metrics:** Objective assessments of a paper's impact and influence are seamlessly integrated into the context, guiding informed decision-making.
- **Requirements and Schemas:** These define the fundamental elements and structures necessary to adeptly acquire and model contextual information.

Beyond enhancing individual papers, context extends its reach, enabling us to:

- **Group related papers:** By encapsulating multiple works within a single context, we foster a deeper understanding of research subfields and trends.
- **Connect to other entities:** Authors, journals, and even sub-contexts seamlessly integrate into the fabric of a paper's context, revealing valuable relationships and dependencies.

This intricate interconnectedness empowers Onto-CCR to:

- **Enhance Learning:** Furnishing a comprehensive context for each paper empowers our system with a deeper understanding of the research landscape.
- **Improve Recommendation Accuracy:** Contextual awareness enables our system to suggest research resources that are truly relevant and impactful for individual needs.

In essence, the "**Context**"(14) class is not merely a data structure; it is the beating heart of Onto-CCR, steering its ability to comprehend, analyze, and connect with research knowledge in a meaningful way.

$$
\begin{aligned}
&\textbf{Context} \sqsubseteq Onto-CCR \ \cap \\
&= Multi \ \ ContextAttribute(xsd:String) \ \cap \\
&= 1 \ \ ContextLevel(xsd:String) \ \cap \\
&= 1 \ \ ContextLifeCycle(xsd:String) \ \cap \\
&= 1 \ \ ContextState(xsd:String) \ \cap \\
&= 1 \ \ ContextType(xsd:String) \ \cap \\
&\exists \ \ BelongTo : Onto-CCR.ContextModel \ \cap \\
&\exists \ \ HasSchemes : Onto-CCR.ContextSchemes \ \cap \\
&\exists \ \ HasQuality : Onto-CCR.QualityOfContext \ \cap \\
&\exists \ \ HasSubContext : Onto-CCR.Context \ \cap \\
&\exists \ \ HasSuperContext : Onto-CCR.Context \ \cap \\
&\exists \ \ HasRequirement : Onto-CCR.RequirementofContext \ \cap \\
&\textbf{Context Schemes} \sqsubseteq Onto-CCR.Context \ \cap \\
&= Multi \ \ ContextAcquisitionSchema[Name, ID, Apt, CA, Csbt, Fbt, Qod, Rbt, VL](xsd:Inetger) \ \cap \\
&= Multi \ \ ContextModelingSchema[ID, Name, LevelDegree, LevelProcess](xsd:Inetger) \ \cap \\
&\exists \ \ HasTime : Onto-CCR.Time \ \cap
\end{aligned}
\tag{14}
$$

In the transformation of existing paper data into this comprehensive ontology, our module systematically dissects each record into constituent elements such as Title, Abstract, Authors and others. It then identifies relevant concepts within Onto-CCR, generating individuals for them and populating these entities with extracted properties and values. The module concludes by applying sophisticated semantic ontology rules to establish connections between individuals. Going beyond mere co-occurrence, these rules unveil concealed relationships such as "CitedBy" (indicating when paper X appears in the references of paper Y) or "Co-author" (highlighting instances where authors collaborate on the same paper). This meticulous process results in a robust ontology model teeming with interconnected individuals, poised for exploration and in-depth analysis.

In Figure 15, we present a concise example of an ontology individual model resulting from the process of transforming a dataset using the Onto-CCR ontology model. The figure illustrates various nodes and edges that signify relationships between different individuals.

- **Circles**: Represent individuals; for instance, the yellow circle labeled "Open-SBS" represents a paper with the title ''Open-SBS: Smart Building Simulator.'
- **Circles with Dotted Lines**: Signify ontology classes from which the individuals are derived.
- **Green & White Rectangles**: Depict data properties of individuals. These data type properties hold real data for each mentioned type.
- **Blue Rectangles**: Represent object properties indicating relationships between existing individuals, such as PublishedIn ,HasAuthor and HasProfile.

E Detection and Retrieval of Missing Data via Query Modelling Module

For instance, Figure 15 illustrates an ontology individual model that contains numerous missing data points. For example, the paper "Xu Du" lacks information about the corresponding author, while the paper by "Sánchez" is missing both the publication date and keywords. Additionally, there are other instances of missing data. To identify and address these missing data points, the Query Modeling Module employs a set of semantic rules using SWRL and SPARQL for each individual. For instance:

- Rule to identifies papers without a designated corresponding author (SPARQL).

```
SELECT ?paper ?missing WHERE {
  ?paper rdf:type Onto-CCR:Paper .
  NOT EXISTS { ?paper Onto-CCR:hasCorrespondingAuthor ?author } .
  BIND("Corresponding Author" AS ?missing) .
```

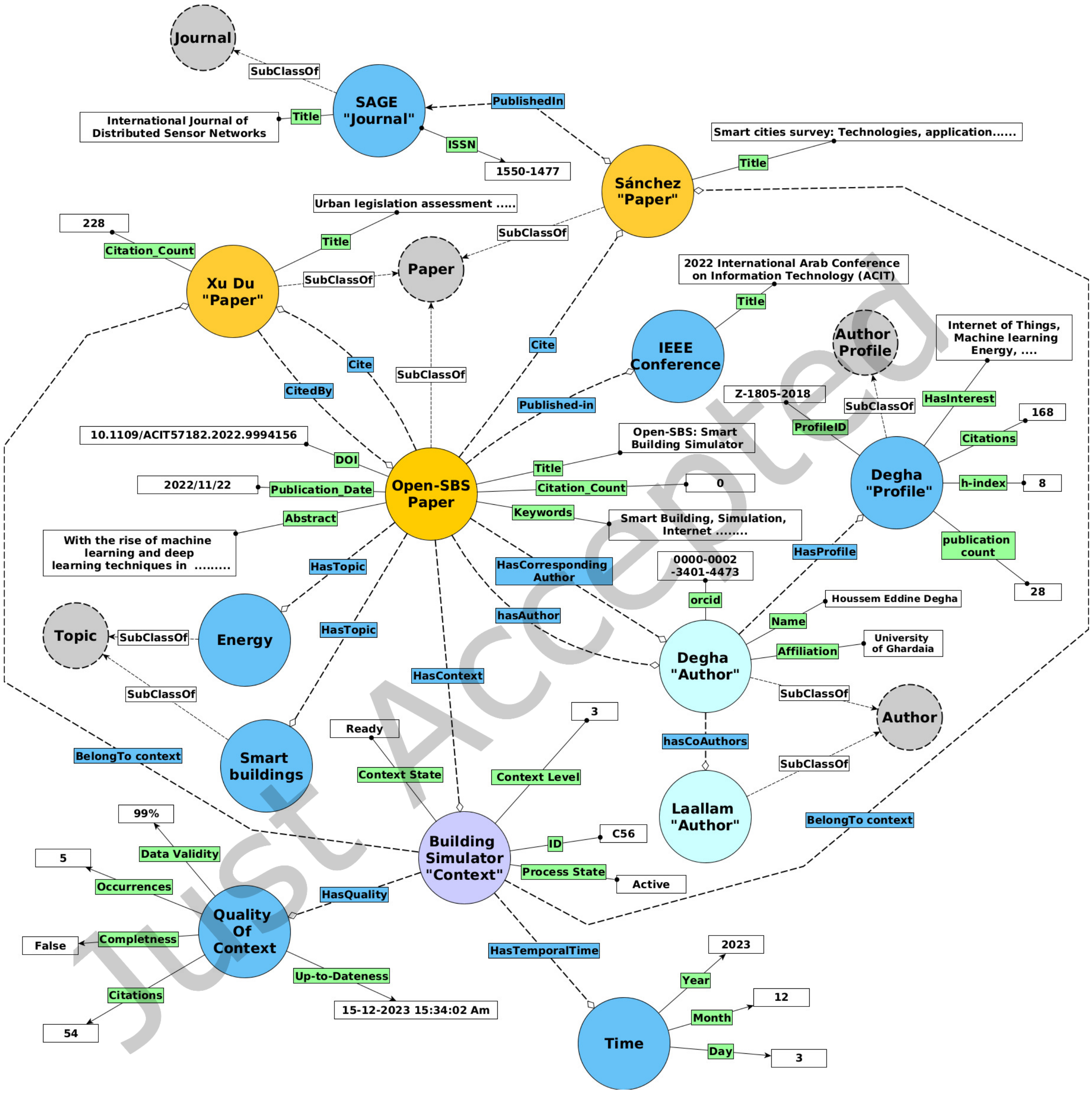


Fig. 15. Illustration Example of Transforming Dataset into Ontology Individual Model.

```
}
\label{rule1}
```

Listing 1. SPARQL Query for Missing Corresponding Author

At the stage where all missing data has been detected, the Query Modeling Module (QMM) identifies the most significant individuals with top-ranking values based on the calculation of the betweenness centrality formula (Equation 1). Only these individuals are considered for generating queries to search for missing data. The QMM generates a query to search for each piece of missing data. This process begins by taking as input the existing data for each individual and creating a search query based on the available information. Each query is dedicated to searching on particular web source. Our system implements functions to search over 12 scientific publishers such as IEEE, ACM, Elsevier, Springer, and others. For example, the paper by the author "Sánchez" is missing the publication date, the function will use existing information, such as paper title and journal, to generate the appropriate URL query to search for this article on the 'sagepub' platform. The generated query would look like this:

```
https://journals.sagepub.com/action/doSearch?AllField=Smart+cities+survey%3A+Technologies%2C+application+domains+and+challenges+for+the+cities+of+the+future&SeriesKey=dsna&AfterYear=2019&BeforeYear=2019&ContentItemType=research-article&rel=&access=18&startPage=&ContentItemType=review-article
```

All the generated queries are saved in the query buffer and will be sent later to the Searcher Model for scraping and crawling the missing data from the web.

F Context Reasoning rules

Below, we present some of the SWRL rules seamlessly integrated into our system, providing insight into the underlying reasoning mechanisms that govern its functionality :

(1) **Selecting Papers with Matching Keywords:** This rule selects papers for a user if the paper and the user share at least one common keyword.

```
Paper(?p) ^ User(?u) ^ hasKeywords(?p, ?k) ^ hasKeywords(?u, ?k) -> RecommendedPaper(?p, ?u)
```

(2) **Selecting Papers by Same Authors' Co-Authors:** This rule selects papers for a user if the paper's author has co-authored with authors the user has previously worked with.

```
Paper(?p) ^ User(?u) ^ hasAuthor(?p, ?a) ^ hasCoAuthors(?a, ?c) ^ hasCoAuthors(?u, ?c) -> RecommendedPaper(?p, ?u)
```

(3) **Selecting Papers from the Same Journal:** This rule selects papers for a user if the paper is from the same journal that the user has previously read.

```
Paper(?p) ^ User(?u) ^ hasJournal(?p, ?j) ^ hasReadJournal(?u, ?j) -> RecommendedPaper(?p, ?u)
```

(4) **Selecting Papers with High Citations:** This rule selects papers for a user if the paper has more than 100 citations.

```
Paper(?p) ^ hasCitationCount(?p, ?c) ^ swrlb:greaterThan(?c, 100) -> RecommendedPaper(?p, ?u)
```

(5) **Selecting Papers based on Author's Research Interests:** This rule selects papers for a user if the paper's author shares research interests with the user.

```
Paper(?p) ^ User(?u) ^ hasAuthor(?p, ?a) ^ hasResearchInterests(?a, ?r) ^
    ↪ hasResearchInterests(?u, ?r) -> RecommendedPaper(?p, ?u)
```

(6) **Selecting Papers with High Rating in User's Research Area:** This rule selects papers for a user if the paper has a high rating (greater than 4.0) and shares common keywords with the user's research interests.

```
Paper(?p) ^ User(?u) ^ hasRating(?p, ?r) ^ hasResearchInterests(?u, ?ri) ^
    ↪ hasKeywords(?p, ?k) ^ hasKeywords(?ri, ?k) ^ swrlb:greaterThan(?r, 4.0)
    ↪ -> RecommendedPaper(?p, ?u)
```

(7) **Retrieve Papers Based on Journal Impact Factor:**

```
Paper(?paper) ^ PublishedIn(?paper, ?journal) ^ JournalImpactFactor(?journal, ?
    ↪ factor) ^ swrlb:greaterThan(?factor, 5.0) -> RelevantPaper(?paper)
```

(8) **Retrieve Papers with High Accuracy:**

```
Paper(?paper) ^ Accuracy(?paper, ?accuracy) ^ swrlb:greaterThan(?accuracy, 0.9)
    ↪  -> RelevantPaper(?paper)
```

G Illustrative Example: The Algorithm Iterative Process in the Filtering Article Module

Figure 16 provides a visual representation of a sample graph, highlighting the intricate relationships between words and articles through semantic connections. To illustrate the algorithm's iterative process, we turn to Figure 17(a). In this initial stage, blue circles represent articles with their corresponding power scores, while the red line denotes the power threshold. Articles below this threshold, marked as red circles, are subsequently removed from the dataset. Figures 17(b) and 17(c) depict subsequent iterations, demonstrating the continuous removal of articles with insufficient power. In this specific example, the objective is to select the top 3 articles. After 50 iterations, the algorithm successfully identifies these articles, showcased as green circles in Figure 17(d). These articles emerge as the most relevant and influential within the dataset, determined through the algorithm's evaluation of their power scores.

In conclusion, the post-ranking stage, integrated with HLART and ICA-CRMAS algorithms, plays a pivotal role in transforming our FAM module into a powerful tool for personalized and effective scholarly recommendation. By unveiling latent aspects, incorporating user preferences, and dynamically adapting to individual needs, the system empowers researchers to navigate the vast and ever-growing landscape of academic literature with greater efficiency and precision.

H Detailed Evaluation Metrics

In our comprehensive evaluation of the performance and effectiveness of our citation recommendation approach, we employ a diverse set of evaluation metrics that gauge the quality and utility of our recommendations. The ranking-based metrics are instrumental in assessing the precision and diversity of our citation recommendations. They provide a quantitative measure of how well our approach organizes and presents candidate citations based on their similarity to the query text. These metrics carry significant implications:

- **Mean Reciprocal Rank (MRR):** MRR evaluates our ability to position the most pertinent citation at the pinnacle of the ranked list of candidates. A higher MRR value signifies our approach's proficiency in

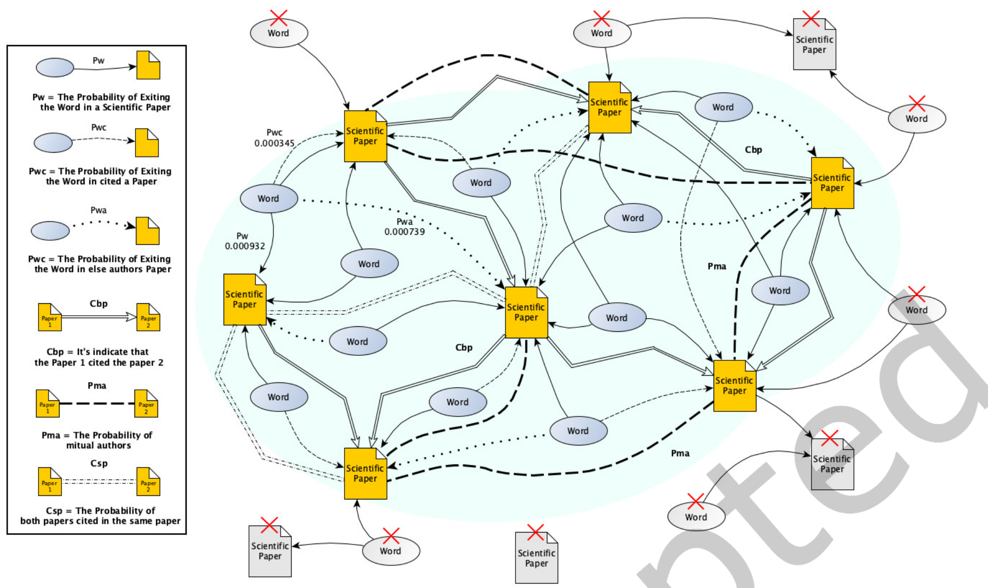


Fig. 16. Illustration of a Graph Constructed by Flittering Agents to Represent Concepts and Articles.

elevating the most relevant citation closer to the top, indicative of enhanced recommendation accuracy.

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{\text{rank}_i} \tag{15}$$

$|Q|$ is the total number of queries.
rank_i is the rank of the first relevant citation for query i.

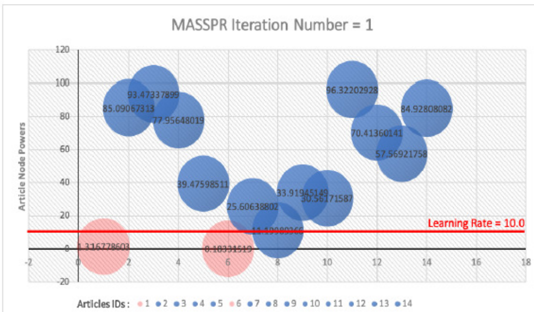
- **Recall@k:** Recall@k measures the proportion of relevant items retrieved in the top-k positions of the ranked list. It evaluates the ability of the recommendation system to find all relevant citations early in the list.

$$Recall@k = \frac{1}{M} \sum_{i=1}^{M} \frac{|R(p_i) \cap T(p_i)|}{|T(p_i)|} \tag{16}$$

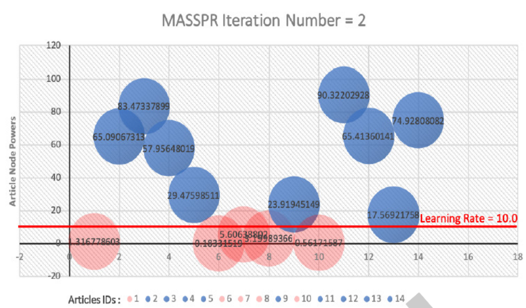
Where:

– M is the total number of queries (test papers).
– $|R(p_i) \cap T(p_i)|$ is the count of relevant papers retrieved in the top-k for query p_i.
– $|T(p_i)|$ is the total number of relevant papers for query p_i.

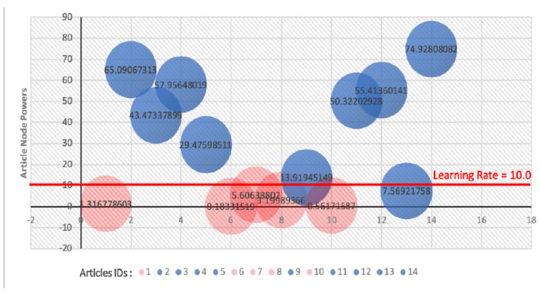
A higher Recall@k value indicates that a larger proportion of relevant papers have been retrieved within the top-k positions, reflecting better performance in capturing relevant citations early in the recommendation list.



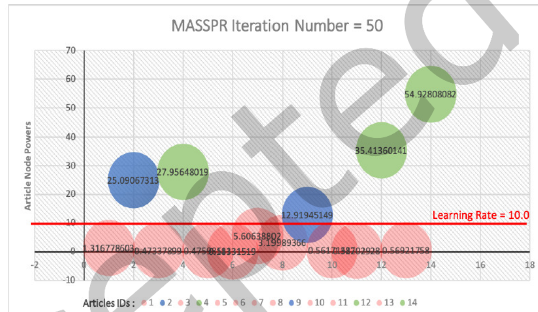
(a) ICA-CRMAS Algorithm iteration Number = 1.



(b) ICA-CRMAS Algorithm iteration Number = 2.



(c) ICA-CRMAS Algorithm iteration Number = 3.



(d) ICA-CRMAS Algorithm iteration Number = 50.

Fig. 17. ICA-CRMAS Algorithm iteration

- **Precision at K (P@K):** P@K evaluates our proficiency in placing relevant citations within the top-K positions of the ranked list. A higher P@K value signifies our approach's effectiveness in bringing relevant citations to the forefront, contributing to both recommendation accuracy and diversity.

$$P@K = \frac{1}{|Q|}\sum_{i=1}^{|Q|}\frac{|R_i \cap \text{top}_K|}{K} \tag{17}$$

$Where:$

$|Q|$ is the total number of queries.

$|R_i \cap \text{top}_K|$ is the number of relevant citations in the top K positions of the ranked list for query i.

- **factorized_top_k**: The factorized_top_k metric provides a comprehensive evaluation of the recommendation quality at various levels of recommendation list length. It helps assess how well the model ranks the items, considering both relevance and diversity.The mathematical formula in LaTeX for the factorized_top_k metric can be represented as follows:

$$\text{factorized\_top\_k} = \frac{1}{N}\sum_{i=1}^{N}\text{precision\_at\_k}(i)\cdot\text{recall\_at\_k}(i) \tag{18}$$

This formula calculates the average precision and recall across all positions in the ranked list, providing a measure of the overall recommendation quality.

I Detailed Participant Group Information

To enhance our understanding of the variations in usability perceptions across diverse user groups, we implemented a segmentation strategy based on two primary criteria: participants' experience level with citation recommendation systems and their primary research discipline. The segmentation criteria are presented below:

- **Experience Level:**
  - Novice: Participants with no prior experience using citation recommendation systems.
  - Intermediate: Participants with some experience using similar systems, but not necessarily regular users.
  - Expert: Participants who regularly use citation recommendation systems in their research workflows.
- **Research Discipline:**
  - STEM: Participants whose primary research field falls within Science, Technology, Engineering, or Mathematics.
  - HASS: Participants whose primary research field falls within the Humanities, Arts, and Social Sciences.

This approach guarantees the representation of a wide array of perspectives, allowing for an in-depth analysis of potential divergences in user experiences within these specified categories. Table 15 provides a comprehensive overview of the composition of each group, offering detailed insights into their characteristics and demographics.

Group	**Research Role**	**Demographics**	**Focus of Research**
Novice STEM (1)	16 Researchers, 8 Students	24 (15M, 9F)	Biology, Computer Science, Physics
Intermediate STEM (2)	25 Researchers, 14 Students	39 (18M, 21F)	Chemistry, Engineering, Mathematics
Expert STEM (3)	24 Researchers, 6 Students	30 (14M, 16F)	Environmental Science, Medicine, Psychology
Novice HASS (4)	21 Researchers, 9 Students	30 (18M, 12F)	History, Literature, Philosophy
Intermediate HASS (5)	20 Researchers, 15 Students	35 (18M, 17F)	Art History, Economics, Linguistics
Expert HASS (6)	9 Researchers, 20 Students	29 (13M, 16F)	Anthropology, Political Science, Sociology

Table 15. Detailed Participant Group Information

J Study Procedures

This section outlines the meticulous procedures employed in our user study, designed to ensure a thorough evaluation of ICA-CRMAS's usability (Algorithm 2). Our commitment to rigorous assessment is evident in the detailed exploration of user tasks, participant interactions, and the specific instructions provided. To enhance the user experience and evaluate a broader spectrum of system functionalities, we developed seven specific user tasks. These tasks carefully simulate diverse research scenarios that users might encounter in real-world settings, providing a comprehensive understanding of the system's usability across various research contexts. Each participant in the six groups was tasked with completing the seven designated tasks in the specified order (refer to Table 16). The study spanned a week, during which participants diligently executed the tasks and offered feedback in accordance with the predefined schedule. Through the systematic collection of data from each participant across all tasks, we conducted a thorough analysis, gaining valuable insights into the strengths and weaknesses of ICA-CRMAS in comparison to other baseline platforms. The subsequent section elucidates the details of the Results and Comparisons.

Received 13 July 2023; revised 9 July 2024; accepted 17 July 2024

Scenario	**Task**	**SUS Focus**
1 : Recommendation Match & Comparison	- Perform the following steps for each system: • ICA-CRMAS: Analyze recommended citations and assess alignment with research interests. • Google Scholar: Utilize search functionalities and recommendations. Compare results with ICA-CRMAS. • Base-Search: (Follow the same steps as with Google Scholar, comparing results.) - Compare recommendations: Identify strengths and weaknesses in recommendation accuracy, relevance, diversity, and user interface.	Evaluate the effectiveness, efficiency, and satisfaction of recommendation engines across all three systems.
2 : Explanation Transparency & Comparison	- Perform the following steps for each system: • ICA-CRMAS: Examine the explanation provided for recommending the chosen article. • Google Scholar and Base-Search: Observe if any explanation or justification is provided for recommended articles. - Compare transparency: Evaluate how each system helps understand the rationale behind recommendations.	Assess usability of explanation features across all three systems, focusing on understandability, transparency, and user trust in recommendations.
3 : System Comparison & Preference	- Compare overall experience using each system. - Identify strengths and weaknesses, considering factors like search options, recommendation accuracy, explanation/justification, ease of use, and user interface design. - Indicate preference for a system and explain reasoning based on research needs.	Assess overall usability and user satisfaction with ICA-CRMAS compared to baseline systems, allowing for a comprehensive understanding of user preferences and the system's relative strengths and weaknesses in the research workflow.
4 : Figure-based Recommendation Comparison	- Perform the following steps for each system: • ICA-CRMAS: Observe if figures are included in recommended citations and explanations. • Google Scholar and Base-Search: Note whether these platforms offer functionalities related to figure-based search or figure analysis within recommendations. - Compare systems in their ability to utilize figures for recommendations.	Assess the effectiveness and efficiency aspects of the systems, focusing specifically on their ability to leverage figures for improved recommendation accuracy.
5 : Serendipitous Discovery & Comparison	- Perform the following steps for each system: • ICA-CRMAS: Analyze recommended citations and identify instances where ICA-CRMAS suggested articles that broadened research perspective. • Google Scholar and Base-Search: Explore recommendations and note if they suggest articles outside the scope of the initial search query. - Compare systems in their ability to promote serendipitous discovery.	Explore the diversity and serendipity aspects of the recommendations across all three systems.
6 : Personalization & User Profile Management	- Explore options available for managing user profile in ICA-CRMAS. - Add or modify research interests and preferred citation styles. - Observe how ICA-CRMAS adapts recommendations based on the updated profile. - Conduct a new search on a different topic of interest and analyze recommendations.	Assess usability and effectiveness of the user profile management features in ICA-CRMAS, focusing on ease of use and the impact of profile updates on recommendation personalization.
7 : System Performance & Scalability	- Perform the following steps: • Conduct multiple searches on various topics within the research field using ICA-CRMAS. • Observe the response time of the system for each search and recommendation generation. • Assess the stability of ICA-CRMAS under extensive tasks, noting any instances of slowdowns or errors. - Compare experience with baseline systems: Reflect on the experience using Google Scholar and Base-Search for similar large-scale searches.	This task examines the system's efficiency and scalability under heavy usage, focusing on performance, responsiveness, and stability compared to baseline platforms.

Table 16. Task Details

Algorithm 2: Participant Study Procedure

Input: Participant
Output: Participant SUS scores, feedback, and completed tasks
Begin Welcome and Introduction

1. Greet the participant and explain the study objectives.
2. Obtain informed consent.
3. Briefly introduce the systems involved (ICA-CRMAS, Google Scholar, and Base-Search).

Familiarization

1. provide a brief demonstration on the systems involved, especially ICA-CRMAS.
2. Allow participants to ask clarifying questions.

Task Loop for $t \in \{1, 2, \ldots, 7\}$ **do**

  1. Present the instructions for task t clearly and concisely.
  2. **Participant completes task** t**:**
     - Monitor progress and provide assistance if needed.
     - Record the time taken to complete the task.
  3. Collect participant data

SUS Questionnaire

1. Administer the System Usability Scale (SUS) questionnaire to collect usability feedback on ICA-CRMAS.

End

Algorithm 3: SUS Score Calculation

Input: Participant SUS scores, participant group information (Novice, Intermediate, Expert)
Output: Adjusted SUS scores for each group, overall SUS score
Begin
Separate Data by Group

1. Divide participant data into separate datasets for each group (Novice, Intermediate, Expert).

Calculate Adjusted Scores for Each Group for $g \in \{Novice, Intermediate, Expert\}$ **do**

  1. Separate odd and even SUS scores within group g.
  2. For each score in the odd and even sets:
     - Calculate adjusted score
  3. Combine adjusted scores for odd and even sets back into a single dataset for group g.

Calculate Overall SUS Score

1. Combine adjusted scores from all groups into a single dataset.
2. Calculate the average of all adjusted scores in the combined dataset. This is the overall SUS score.

End