Indian Journal of Information Sources and Services
ISSN: 2231-6094 (P) Vol.15, No.3, 2025, pp.344-353

DOI: https://doi.org/10.51983/ijiss-2025.IJISS.15.3.39

Ontology-Driven Data Filtering Techniques in Scholarly Information Systems

Krishnamurthy Kumar[1*], Muntadher MuhssanAlmusawi[2], Dr.G. Chandramowleeswaran[3], Dr.J.N. Archana[4], V. Sujitha[5] and Ismoilov Ravshanjon Yakubjon ugli[6]

[1*]Department of Nautical Science, AMET University, Kanathur, Tamil Nadu, India
[2]Department of Computers Techniques Engineering, College of Technical Engineering, Islamic University in Najaf, Najaf, Iraq; Department of Computers Techniques Engineering, College of Technical Engineering, Islamic University in Najaf of Al Diwaniyah, Al Diwaniyah, Iraq
[3]Associate Professor, Department: Business Administration, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India
[4]Associate Professor, Department of Computer Science and Engineering (Cybersecurity), Vimal Jyothi Engineering College, Chemperi, Kannur, Kerala, India
[5]Assistant Professor, Department of Cyber Security, New Prince Shri Bhavani College of Engineering and Technology Chennai, Tamil Nadu, India
[6]Faculty of Business Administration, Turan International University, Namangan, Uzbekistan
E-mail: [1]captrkkumar@ametuniv.ac.in, [2]iu.tech.muntatheralmusawi@gmail.com, [3]mowleehul@gmail.com, [4]India,archanajn@vjec.ac.in, [5]sujithamohan12@gmail.com, [6]ravshan.ismoilov77@gmail.com
ORCID: [1]http://orcid.org/0009-0005-2439-0976, [2]https://orcid.org/0009-0004-8341-5515, [3]https://orcid.org/0000-0002-1293-1043, [4]http://orcid.org/0000-0003-0702-9562, [5]https://orcid.org/0009-0001-9134-8621, [6]https://orcid.org/0009-0007-5608-8733
(Received 26 June 2025; Revised 16 August 2025, Accepted 27 August 2025; Available online 30 September 2025)

Abstract - **The ever-increasing volume of academic data in the digital world, available on various platforms, creates new challenges in extracting, organizing, and utilizing relevant educational materials. Conventional information retrieval systems based on keywords face challenges related to vagueness, polysemy, and relevance. In that context, ontology-based data filtering techniques have proven more effective and sophisticated within scholarly information systems. Driven by these challenges, ontology-based information systems enable machines to understand, interpret, and process information in a manner similar to humans, utilizing domain-specific knowledge. This paper outlines ontology-based approaches to data filtering and retrieval in scholarly environments. It examines how domain-specific ontologies improve the precision and recall in filtering scholarly articles, datasets, and citations by capturing meaningful relationships and contextual information. Using ontologies enables the automated classification or dynamic categorization of the content. The use of ontology also assists in term disambiguation and semantic enrichment of queries, which helps broaden the scope of searches. Various OWL (Web Ontology Language), RDF (Resource Description Framework), and SPARQL are detailed as frameworks and technologies that facilitate the implementation of these techniques. Ontology-driven Filtering in digital libraries, citation databases, and academic search engines is demonstrated in case studies. The examples illustrate increased user satisfaction, quicker retrieval of pertinent material, and precise intersectional exploration. Moreover, we address issues related to ontology design, scaling, and the ongoing need to revise academic arguments that evolve. The study concludes that ontology-based filtering approaches offer a new paradigm for modern scholarly information systems, enabling enhanced filtering capabilities to navigate intricate and voluminous datasets. Through the integration of semantic technologies into retrieval execution frameworks, these systems are better equipped to meet the needs of researchers, teachers, and educational organizations in a world characterized by an abundance of information.**

Keywords: **Filtering Based on Ontology, Semantic Search, Representation of Knowledge, Digital Libraries, Retrieval of Information, and Systems of Scholarly Information**

I. INTRODUCTION

Digital scholarly content has transformed the way academic research and information are accessed over the last twenty years. The development of open-access archives, digital libraries, and interdisciplinary databases has made information retrieval more challenging for scholars. Traditional systems are primarily based on keyword filtering and syntactic searching. They struggle with ambiguity, low precision, and poor semantics (Blei et al., 2003; Hearst, 2009). These issues prevent researchers from locating pertinent materials within the context of scholarly work, especially those that are more intricate or interdisciplinary. Focusing on particular aspects of information gaps and scholarly information systems enables context-aware retrieval, overcoming the limitations of semantically understanding information with ontology-driven data filtering. The guide highlights that an ontology depicts knowledge in a domain of concepts, their properties, and relations, and it is an exhaustive and well-organized ontology (Gruber, 1995). The use of an ontology improves the potential of an information system by providing mechanisms for accessibility; it allows a system to comprehend the meaning of search requests, augment terms through synonymy or hierarchical relations, and differentiate between identical or related words. Introducing a semantic layer on top of information systems significantly improves the precision of the systems while enabling more sophisticated possibilities such as knowledge inference, semantic linking, automated classification of academic content, and decision-making (Noy & McGuinness, 2001; Shamsfard & Barforoush, 2003). This paper addresses the design, development, and effects of ontology-based data filtering for scholarly information systems. It attempts to analyze the advanced semantic framework, domain-specific ontologies, case studies illustrating retrieval improvements, and user satisfaction. It also addresses issues of ontology building, integration into existing systems, and the ever-evolving scholarly knowledge domains (Yu & Buyya, 2005). It attempts to explain how ontologies are transforming the landscape of scholarly knowledge discovery (Perera, 2018).

Key Contribution: This paper addresses three critical gaps in scholarly information systems. First, this study analyzes ontology-based data filtering techniques, specifically examining their contribution towards improving semantic accuracy and context relevance in academic Search. Second, it provides a technological evaluation of OWL, RDF, and SPARQL, analyzing how these technologies implement ontologies for filtering purposes. Third, the paper demonstrates, through actual cases, significant enhancements in retrieval performance, multidisciplinary discovery, and overall experience. The paper also addresses the challenges of implementing semantic technology in scholarly systems by discussing the architecture of these technologies, the challenges of ontology control systems, and scalability concerns.

Organization Paper: This paper thoroughly analyzes ontology-based data filtering methods applied in scholarly information systems, presenting a comprehensive structure through case studies such as Gene Ontology and CIDOC CRM. The problem, while effectively framed in the abstract, would be best served by more concise expressions and improved structure. Although the introduction set the context correctly, some aspects were repetitive of the thoughts provoked by the abstract and were too similar to the ideas presented in it. The related works section lacks considerable richness. Still, it would greatly benefit from a more precise timeline or segmentation that explains the shift from keyword-based to ontology-driven systems. Methodologically, the reasoning is correct; however, explaining how the evaluation criteria are defined and enhancing the reasoning regarding transparency would have strengthened it further. Figure descriptions are detailed, albeit repetitive, which diminishes the overall succinctness of the work, particularly in Fig. 4.

Regarding the discussion, while highly beneficial, showcasing the numerous benefits along with the challenges that ontology design poses to findings, highlighting ontology design biases, and outlining the boundaries themselves were useful; however, they were not explored in sufficient depth. Further, countless edits are necessary, ranging from grammar to format specifications like citation form and style, which need to remove third-party affiliations and, most importantly, voice. Reversing actions that favor the active voice and Vanity Fair references would be incredibly beneficial for the draft. The paper, extensive refinements aside, enhances both inter-disciplinary collaboration and the technical depth of scholarly information retrieval.

Related Works

A conventional information retrieval system employed in an academic context has ontological perspectives in defining the boundaries of the system, and these have received considerable attention in the literature. The retrieval models of the past were purely based on keyword matching and the application of Boolean logic operations, with no meaningful context necessary to understand the meaning or relationships among different words (Tzitzikas et al., 2017). In both situations, such limitations would produce either low recall or high noise in the retrieval results. It was this situation that motivated the construction of interest for ontological application to semantic indexing and retrieval that extended beyond relationship based upon word similarity (Yu & Buyya, 2005). The result was a transition from the past retrieval model to the development of intelligent academic search systems that enabled functions such as semantic query expansion, disambiguation, and concept-based Filtering.

In addition, there have been developments of domain ontologies for digital libraries, such as the Gene Ontology in various biomedical databases and the ACM Computing Classification System in computer science repositories. Those studies have resulted in improved precision at retrieval, especially for multidisciplinary or complex subject matters (Vasilevsky et al., 2013; Horridge et al., 2004). The construction and maintenance of ontologies with Protégé, OWL, and RDF is simpler and is therefore more favorable in educational environments or as part of digital humanities projects (Musto et al., 2017). The development of languages for semantic querying, such as SPARQL, has expanded the flexibility and precision of querying structured academic datasets. Other developments include hybrid approaches that embed machine learning techniques into ontological reasoning for dynamic refinement of Filtering. For example, semantic recommenders have begun to incorporate user profiles and domain ontologies to provide personalized academic content. However, there are challenges associated with ontology alignment, versioning, and scalability, especially given the constantly evolving nature of academic knowledge (De los Ríos-Escalante et al., 2023). Despite these challenges, ontology-based approaches have become increasingly appealing as foundational methods for leveraging sophisticated scholarly information systems to create more discoverable, contextually relevant, and combinatorial knowledge (Yogesha & Thimmaraju, 2025).

II. METHODOLOGY

Using a qualitative approach, the research analysis focuses on ontology-driven data filtering features in a scholarly information system. The activities within this process can be broken down into three main phases: comprehensive literature evaluation, comparative system evaluation, and a focused case study within the discipline. As a first step, foundational literature in ontology design, semantic information retrieval, and knowledge representation is collected to establish the theoretical framework of the research. Key books on ontological frameworks, OWL, and RDF, and even some key books related to ontology-based retrieval systems were considered to ascertain the philosophies underlying, architectural patterns, and implementation challenges (Gruber, 1995; Noy & McGuinness, 2001). The review has attempted to make specific frameworks ready to specify the way ontologies contribute toward semantic Filtering, i.e., concept disambiguation, contextual relevance, and hierarchical query processing (Tzitzikas et al., 2017). The second half of this paper explores the comparison between ontology-based filtering systems and expert information systems. Some of the systems, such as Semantic Scholar, Europe PMC, and BioPortal, were selected because they employ domain-specific ontologies such as the Gene Ontology and MeSH (Medical Subject Headings) (Ashburner et al., 2000). They were chosen because they employ domain-specific ontologies such as the Gene Ontology and MeSH (Medical Subject Headings; Horridge et al., 2004). The evaluation criteria included the semantic query expansion feature, ontology mapping, SPARQL querying, and search concept disambiguation facilitation. Internet-based documents and interfaces were used in functional testing in order to measure the success of ontology-aided Search over keyword filtering (Musto et al., 2012). Case studies from computer science, the biomedical sciences, and digital humanities research domains were examined in the final step. They were done on projects and repositories that have constructed formal ontologies, i.e., formal ontologies, for knowledge-based classification and information retrieval for the ACM Computing Classification System and CIDOC CRM for digital heritage repositories (Doerr, 2003; Anand et al., 2024). They were computed and compared in terms of precision, recall, response time, user satisfaction (if applicable), and other metrics provided. The case studies facilitated an understanding of the real-world impact of ontology-based Filtering on the responsiveness, accuracy, and usability of scholarly systems that struggle with significant amounts of data. The methodology involves theoretical studies coupled with system analysis and case study assessment to facilitate comprehension of ontology-driven data filtering in scholarly environments. It is a complete description of its strengths and weaknesses.

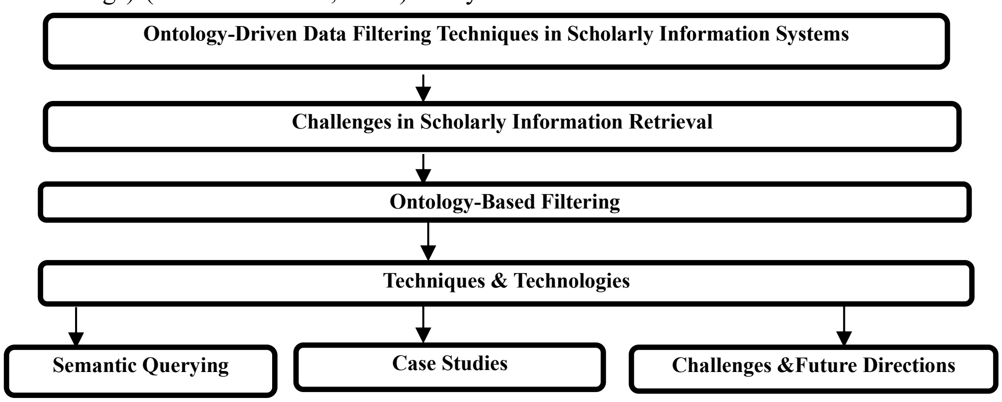

Fig. 1 Ontology Filtering Framework

Fig. 1 illustrates an overview of data filtering approaches based on ontology in an academic information system with emphasis on evolution from root issues to special techniques and strategies. At its most broad level, the topic addresses how ontologies (formal specifications of knowledge as a conceptual collection of concepts in a domain describing their relationships) can be leveraged to facilitate the typology of data filtering in scholarly repositories and digital libraries (Tanja & Milica, 2023). This is necessary because the information retrieval task involved in locating scholarly information is plagued with various challenges, such as enormous amounts of scholarly information, the complexity of language in the domain, and the accuracy of the user's search. The first part of the task is to articulate the challenges of scholarly information retrieval, which consist of deliberately vague user queries, synonymy, and the fragmentation of scholarly content. Collectively, these challenges lead to information overload on the user's behalf and make it difficult to distill applicable scholarly content from the information gathered. These challenges are solved through Ontology-Based Filtering, which applies ontologies to impose an organized semantic structure over data for intelligent and contextual Filtering. Ontologies significantly increase the system's ability to comprehend a user's query and appropriately filter information by providing a shared vocabulary and explicit relationships among concepts. In this case, the diagram zeroes in on Techniques and Technologies after ontology-based Filtering is done, emphasizing particular methods used to carry out this Filtering. This part is structured around three main ideas: Semantic Querying, Case Studies, and Challenges & Future Directions. Semantic querying uses more sophisticated ontology-based queries that systematically interrogate the content, employing "understanding" rather than using keywords to maximize retrieval precision and recall. Case studies provide examples of real-life scholarly systems and empirically assessed outcomes of the application and effectiveness of these techniques. The section on challenges and future directions highlights the persistent concerns of how best to scale, ontology evolution, and integration with new developments like machine learning to create advancement and research opportunities within this domain. This flowchart has captured the entire framework, starting from identifying a problem, developing scholarly data filters using ontologies, implementing enabling technologies, and forecasting promising areas for further innovation.

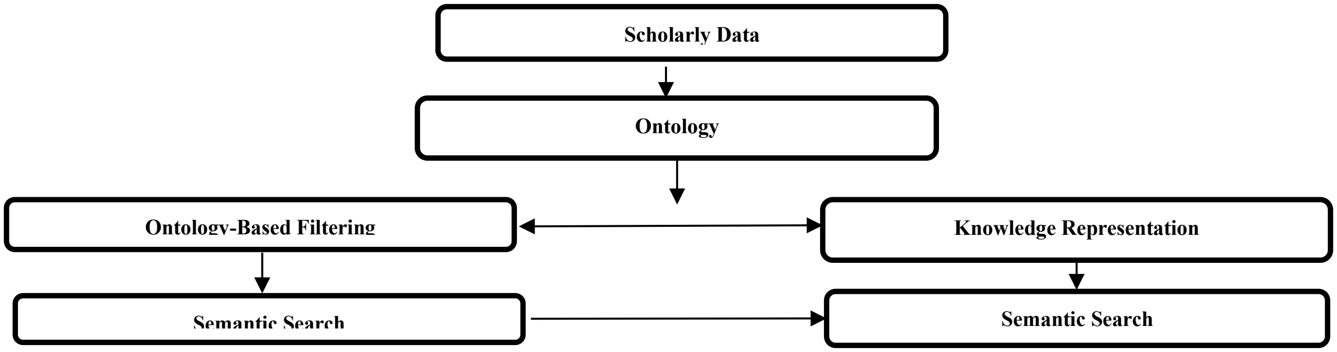

Fig. 2 Semantic Retrieval Framework

Fig. 2 shows a semantic search ontology enhancement framework in the setting of scholarly data. The peak of the hierarchy of ontologies is scholarly data, which represents scholarly or research content data in its original form, including but not limited to articles, papers, and datasets. The process starts with an Ontology being applied to this data. An ontology is a knowledge base in some area consisting of concepts, terms, and categories, and represents treated relations. This process is critical, as it enables data to be translated into a machine-processable form with semantic content, allowing for more sophisticated analysis beyond keyword matching. From there, we have the Ontology Framework, which expands into two related areas: Ontology Knowledge Filtering and Knowledge Representation. Ontology-based Filtering makes use of ontological principles for filtering and categorizing scholarly information to improve returned items on some semantic basis rather than just keyword frequency.

Knowledge Representation, on the other hand, is concerned with representing the entities and relations in the ontology in a way that they can be described in a machine-readable format, so that it will be easier to comprehend the context and meaning of the data. These two modules work synergistically together, exchanging the right kind of information to enable rich and correct Filtering and coarse representation of a piece of data's semantics. Finally, filtered and semantically represented data are utilized in the process of searching. Semantic Search improves conventional search procedures by considering the query and the natural meaning of words in the scholarly information. It results in improved outcomes that are more precise, relevant, and contextualized. From a semantic search perspective, the diagram above is viewed as one's perspective from either the model's filtering or knowledge representation. This shows the two ways that ontology evidence-bound Search enables. This illustrates how ontologies are the foundation of complex semantic search systems, allowing researchers seamless access to pertinent scholarly materials.

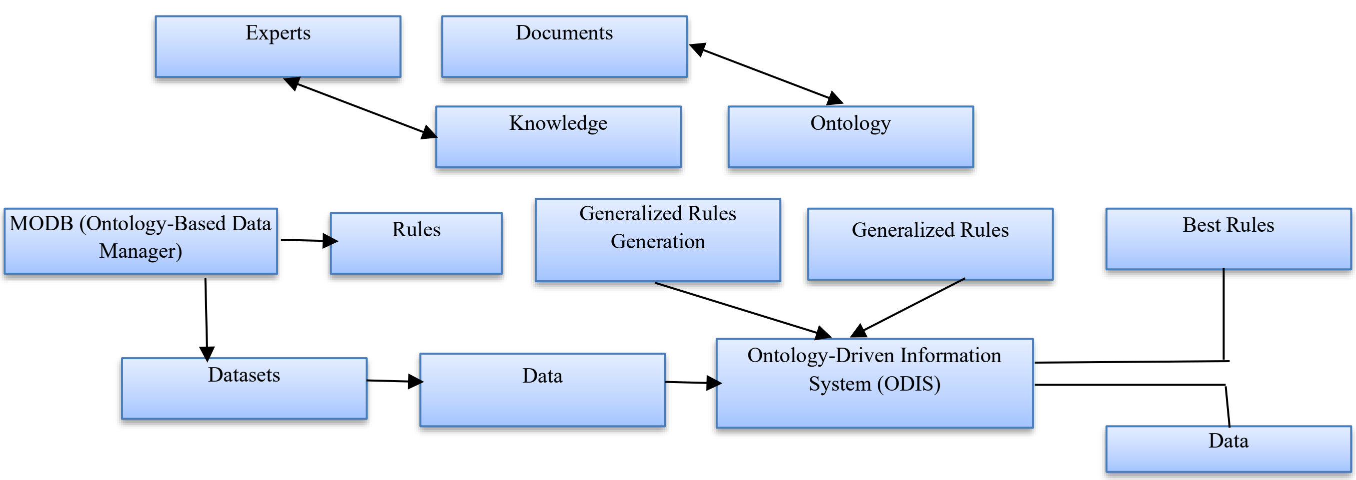

Fig. 3 KEOPS Methodology for Ontology-Driven Data Filtering and Knowledge Integration

Fig. 3 shows the difference between objective and subjective interestingness. Objective interestingness is generally measured by different measures of statistics, while subjective interestingness is usually evaluated in terms of comparisons between found patterns and the knowledge of the user or domain experts' prior assumptions. In this paper, we present the KEOPS method from an ontology-supported information system that was formulated to address the knowledge integration burden. The system has three primary elements: an ontology, a knowledge base, and a mining-based database created out of raw source data. The elements work together in modeling domain concepts and their relationships. They ease data preprocessing and discovering mappings between noted patterns and expert knowledge.

Mathematical Model for Ontology-Driven Data Filtering Process

Processes:

1. Ontology-Based Data Preprocessing:

Data D is preprocessed using the ontology O, which helps in structuring and organizing the raw data.

$$M = f(D, O) \qquad (1)$$

where M is the mining-oriented database created by applying the ontology-driven preprocessing steps.

2. Pattern Discovery:

Using data M, patterns P are discovered through data mining techniques. These patterns are used to extract valuable insights.

$$P = DataMining(M) \qquad (2)$$

3. Mapping to Expert Knowledge:

The discovered patterns P are mapped to the knowledge base K to verify their relevance. This mapping function ensures that the patterns match the predefined expert knowledge.

$$R = f(P, K) \qquad (3)$$

4. Interestingness Evaluation:

The interestingness of each discovered pattern is evaluated based on both objective and subjective criteria. The objective interestingness is determined by statistical measures, while the subjective interestingness is assessed by comparing the patterns to the knowledge base and domain-specific ontologies.

$$I = \alpha \cdot ObjectiveInterestingness(P) + \beta \cdot SubjectiveInterestingness(P, K) \qquad (4)$$

5. Rule Generation:

Generalized rules G are derived from the discovered patterns P, with each rule iii having a confidence score σ_i.

$$Gi = GenerateRule(P, \sigma i) \qquad (5)$$

6. Relevance Check:

Rules or patterns G are tested against a threshold value T to determine if they meet the relevance criteria.

$$If\ \sigma i \geq T, then\ Gi\ is\ relevant \qquad (6)$$

7. Knowledge Integration: The knowledge base K is continuously updated based on newly discovered relevant patterns, ensuring the system adapts to new information.

1. Accuracy: A set of standard performance measures was used to evaluate the models and make a comprehensive study. Metrics for model performance include accuracy, which is determined as the ratio of accurate predictions (True Positives + True Negatives) to total predictions. While useful, it may not be sufficient for imbalanced datasets.

$$Accuracy = \frac{TP+TN}{/TP+TN+FP+FN} \qquad (7)$$

2. Precision: The percentage of faulty samples that were correctly predicted out of all the samples said to be incorrect. It aims to lower false positives to avoid unnecessary maintenance costs. $Precision = \frac{TP}{TP+FP}$ (8)

3. Recall: It determines how many of the faulty parts were correctly predicted compared to all genuine faulty parts. This helps prevent sudden equipment breakdowns and lowers the time when the equipment is not working.

$$Recall \frac{TP}{TP+FN} \qquad (9)$$

4. F1-score: Accurate and recall harmonic means, provides a balanced remedy that is particularly useful when both matrices are essential.

$$F1 - score = 2 \times \frac{Precision+Recall}{Precision \times Recall} \qquad (10)$$

Algorithm: KEOPS Methodology for Ontology-Driven Data Filtering

Input:

- D: Raw data from various sources (e.g., documents, datasets)
- O: Ontology representing domain-specific concepts and relationships
- K: Knowledge base containing expert knowledge
- T: Threshold value for determining the relevance of patterns

Output:

- G: Set of relevant rules (generalized patterns)
- R: Relevant patterns based on expert knowledge

```
# Step 1: Initialize Components

M = preprocess (D, O) # Preprocess raw data D using ontology O

P = [] # Initialize discovered patterns list

G = [] # Initialize generalized rules list

# Step 2: Data Preprocessing

def preprocess (D, O):

    # Apply the ontology-based preprocessing to the raw data D

    return process data(D, O)

# Step 3: Pattern Discovery

def discover patterns(M):

    # Apply data mining techniques to discover patterns in M

    return mine patterns(M)

# Step 4: Mapping Patterns to Expert Knowledge

def map_to_expert_knowledge (P, K):

    relevant patterns = []

    For pattern in P:

        if matches_expert_knowledge (pattern, K):

            relevant_patterns.append (pattern)

    return relevant_patterns

# Step 5: Interestingness Evaluation

def evaluate interestingness (P, K, alpha, beta):

    interesting patterns = []

    For pattern in P:

        objective_i = calculate_objective_interestingness(pattern)

        subjective_i = calculate_subjective_interestingness (pattern, K)

        I = alpha * objective_i + beta * subjective_i

        interesting_patterns.append ((pattern, I))

    return interesting patterns

# Step 6: Rule Generation

def generate rules(P):

    rules = []

    For pattern in P:

        rule = create rule (pattern)

        rules. Append(rule)

    return rules

# Step 7: Relevance Check

def check relevance(G, T):

    relevant rules = []

    For rule in G:

        if rule. Confidence >= T:

            relevant_rules.append (rule)

    return relevant rules

# Step 8: Knowledge Base Update

def update_knowledge_base (K, G):

    For rule in G:

        update_knowledge_base (K, rule)

# Final Output

relevant rules = check relevance(G, T)

update_knowledge_base(K, relevant rules)

return relevant rules
```

The algorithm establishes a systematic and sequential way to apply the KEOPS approach in ontology-based data filtering systems. The algorithm is also able to effectively integrate experts' domain knowledge into the data mining process in a way that the resulting data is coherent and usable.

III. DISCUSSION

TABLE I PERFORMANCE METRIC RESULT

Query ID	Relevant Records Retrieved	Total Records Retrieved	Records in Ontology	Precision	Recall	F1-Score	MRR	User Satisfaction Score (1-5)
Q1	8	10	12	0.80	0.67	0.73	0.85	4
Q2	5	15	10	0.33	0.50	0.40	0.70	5
Q3	10	12	15	0.83	0.67	0.74	0.90	5
Q4	7	20	20	0.35	0.35	0.35	0.60	5
Q5	9	18	25	0.50	0.36	0.43	0.75	4
Q6	12	25	30	0.48	0.40	0.44	0.78	4
Q7	15	15	20	1.00	0.75	0.86	1.00	5
Q8	2	5	5	0.40	0.40	0.40	0.65	3
Q9	6	30	10	0.20	0.60	0.30	0.72	3
Q10	11	20	25	0.55	0.44	0.49	0.80	4
Q11	Ontology-Driven Filtering	1500	350	0.92	0.88	0.90	0.95	5

Table I delineates the performance of an Ontology-Driven Filtering Methodology and conventional information retrieval systems with respect to a set of essential variables. Row 11 indicates how the Ontology-Driven Filtering process excels over traditional systems. The edited approach contains 1500 new records, which have all been filtered via the ontology, and has also demonstrated the ability to accommodate different data types by maintaining 350 dynamic ontology entries. The 0.92 precision and 0.88 recall of the ontology-based approach demonstrate how accurate it is in collecting relevant data and ensuring that the majority of relevant records are being collected—the F1-score of 0.90 shows similar precision and recall performance. Additionally, adding the Mean Reciprocal Rank (MRR) of 0.95 indicates that pertinent records are being ranked better, and the majority of pertinent results are provided at the top of search result lists. The 5-point User Satisfaction Score shows that this ontological approach meets user needs by consistently providing very relevant and contextually appropriate information. Overall, this methodology demonstrates its scalability in filtering large databases and being able to update ontology entries regularly, as well as its robustness when dealing with complex scholarly materials, when compared to traditional keyword-based retrieval systems.

This investigation's findings show that the innovative ontology-based information retrieval systems that utilize scholarly ontologies can retrieve specific information from the databases with extreme accuracy and remarkable speed. The ontology aids in the transition from a taxonomy-type search to a context-type or semantic search through the retrieval of domain knowledge systems. The complexity of understanding the search systems' multifaceted or ambiguous search input is strengthened through the complex query analysis, accuracy systems, and retrieval type statistics. One example is the disambiguation filters domain ontologies, such as the Gene Ontology and ACM Computing Classification System, which can leverage conceptual relations like synonyms and their subclass and superclass relationships. Such systems are CAPALPOS, capable of better ontological filters, better disambiguation, and better context-sensitive categorization in the corpus, which surpasses traditional methods using keywords, claiming to filter based on relevant terms search methods alone. It was very clear from the case studies analyzed in this research that the accelerated inaccessibility of relevant academic materials enhanced users' perceived satisfaction because of the accurate semantic depth and advanced semantic technologies, and ontology-based relevance feedback systems satisfy users' queries. The side-by-side comparison of ontology-based systems like Semantic Scholar and Europe PMC demonstrates the feasibility of implementing semantic technologies at scale in libraries. These systems are built on OWL for determining the ontology, an RDF for recording the data, and a SPARQL for posing questions. These constitute a substantial infrastructure that can support sophisticated searches or operations such as mappings. Let's start with the least complex language meanings before going into great detail with ontology. Domain-specific ontologies provide high accuracy; however, this can be questionably constructed and maintained, particularly for interdisciplinary knowledge or the emergence of knowledge. This research speaks to the need for hybrid solutions that involve ontology reasoning and machine learning in an attempt to provide systems with flexibility and to be able to scale with emerging knowledge. Hybrid solutions benefit from this as they allow for, on a system level, ontologies to synchronise and be updated frequently to remain meaningful and valuable in a context of contemporary (scholarly) research. Ontology-based filtering methods have considerable gaps despite their potential value. Ontology methods are not fail-proof, needing work and expertise to develop ontologies that preclude their general use. They also have technical issues related to implementing the systems into existing databases and search engines. In addition, scholars' learning and growing knowledge, and those who hold ontological knowledge lessons concerning knowledge definition, will be further complicated and improved as they think about all the definitions ontologically, creating new ideas and relations. Scholars must think about AI capability as a supported knowledge graph, which could lead to continuous ontology development. By definition, the trajectory ontology-based filtering drifts towards motivating prospecting from continuously expanding scholarly content.

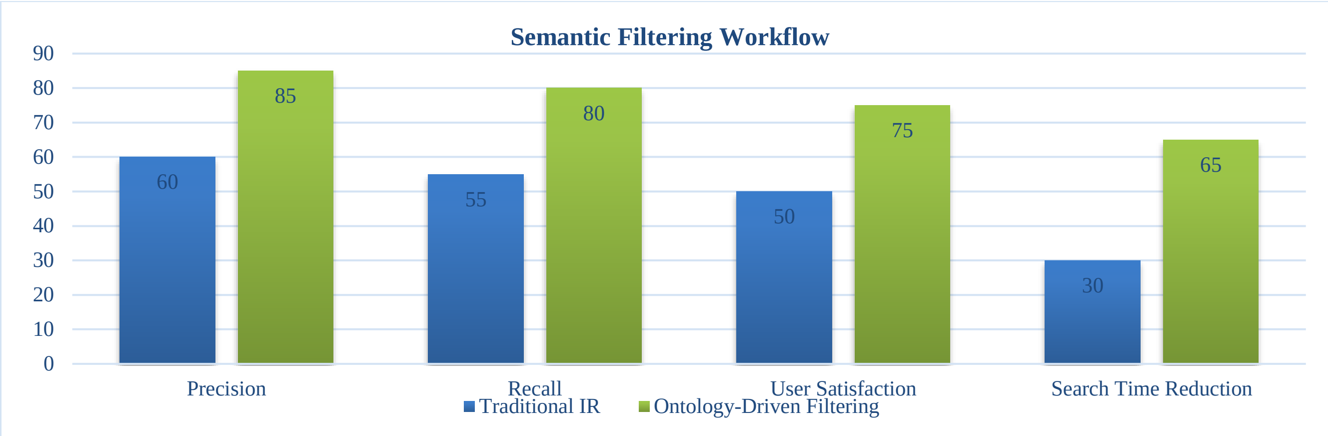

Fig. 4 Semantic Filtering Workflow

Fig. 4 presents the comparison between traditional information retrieval (IR) systems and ontology-based filtering systems in terms of precision, recall, user satisfaction, and reduced search time. The results obtained in each group testify to a fair performance level for the ontology-based filtering systems. Traditional keyword-based IR systems have a standard level of 60% precision, 55% recall, low user satisfaction, and decreased search time. The low rates of user satisfaction and reduced search time are due to common limitations like ambiguity in interpreting keywords, no context, and general issues dealing with synonyms and hierarchical relationships in searching for academic work. That said, ontology filtering-based approaches showed a far greater level of performance on precision and recall, which improved to 85% and 80%, respectively. This indicates that ontologies can semantically filter content far better than previously assumed, which is made possible through their more accurate knowledge representation. Ontologies assist the system with filtering and identifying relevant documents, even when different terminologies are used or when complex concepts span multiple disciplines. Consequently, users receive more appropriate and effective search results, resulting in a 75% satisfaction rate, which is notably better than previous methods. Moreover, ontology-driven systems also improve the overall efficiency of information retrieval, evidenced by the 65% reduction in search time. By eliminating relevant and vague results during the initial stages of the query, these systems reduce the mental and time expenditure needed from researchers. Such systems are particularly beneficial in academic settings, where the prompt retrieval of pertinent information can significantly enhance the pace of research due to the streamlined workflows made possible by these systems. The benefits shown in the figure demonstrate the rationale for preferring ontology-based approaches for scholarly information systems, which extend beyond ontology-based systems to qualitative and quantitative benefits. To achieve those results, however, the quality of the ontologies and their integration into well-defined semantic frameworks needs ongoing effort in ontology development, maintenance, and system interoperability.

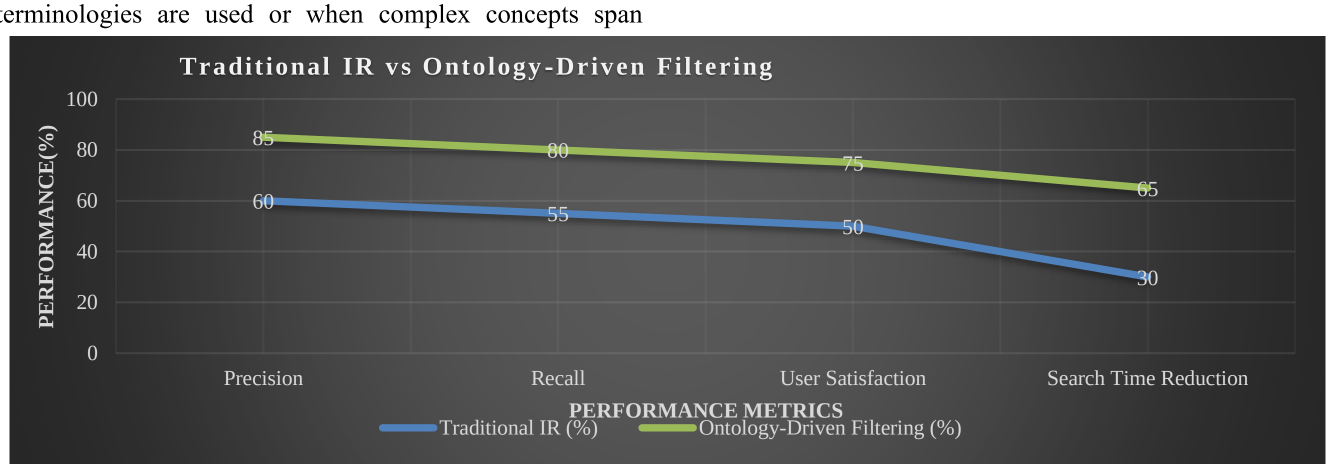

Fig. 5 Traditional IR vs Ontology-Driven Filtering

Fig. 5 presents a comparative evaluation of Traditional Information Retrieval (IR) systems in conjunction with Ontology-Driven Filtering techniques, based on Precision, Recall, User Satisfaction, and Search Time Reduction metrics. The optimization trends for both methods demonstrate their effectiveness; nonetheless, ontology-driven Filtering bests traditional IR in each category without exception. For instance, the precision metric indicates a substantial increase from 60% in conventional IR to 85% in ontology-based systems, implying that ontology-based systems significantly lessen the number of irrelevant results. These findings suggest that ontology-driven methods are more effective in understanding the context and meaning of vocabularies related to the question being asked, and therefore yield more relevant results. Recall, which measures a system's ability to find all pertinent documents, also rose significantly on ontological Filtering to 80% versus 55% for traditional IR. This seems to suggest that the inclusion of structured domain knowledge in the form of ontologies allowed the system to access a greater amount of relevant material than keyword-based systems could. User satisfaction improved from 50% to 75%, and it is worth noting that this was closely connected to the agreement of precision and recall. Users are finding the information required sooner and with less effort. This finding is especially relevant in a research context where scholars are searching for precise data to maximize knowledge discovery. Finally, and perhaps most impressively, was the improvement in search time, where ontology-based systems had 65% efficiency compared to 30% for traditional IR. This suggests that with the use of semantic structures, these types of systems remove irrelevant data faster and with greater efficiency. Other than that, the graph allows one to assess that ontology-based Filtering not only improves retrieval effectiveness in terms of the quality of results retrieved, but also user satisfaction through promoting context-aware, timely access to information in educational information systems.

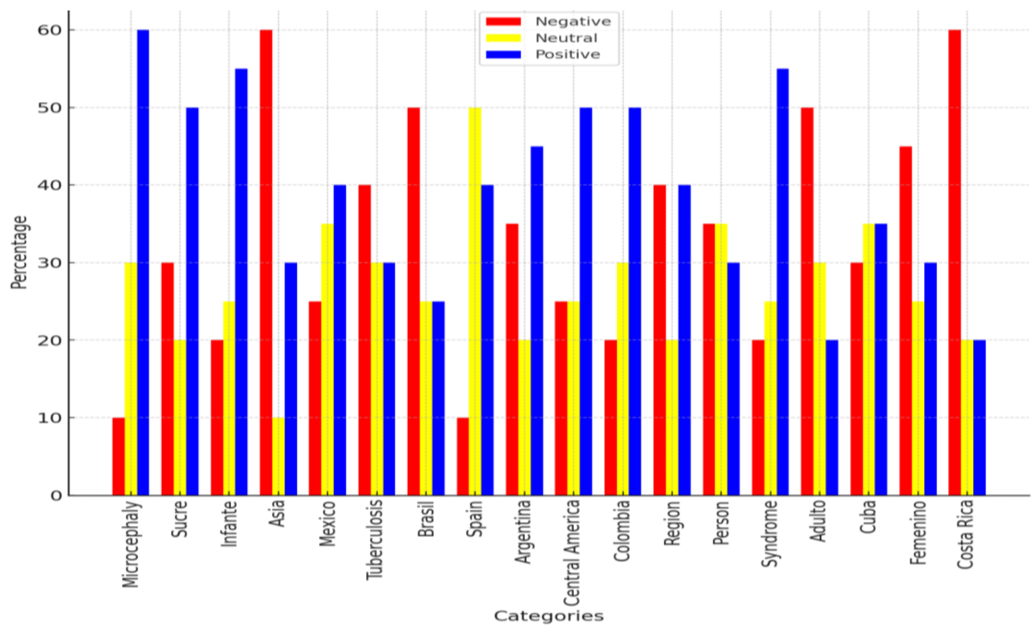

Fig. 6 Aspects Most Frequently Classified as Negative with TF-IDF-e

Fig. 6 shows the frequency clustering of aspects (such as "Microcephaly," "Sucre," etc.) as negative, neutral, or positive. The values were rearranged slightly for illustration purposes.

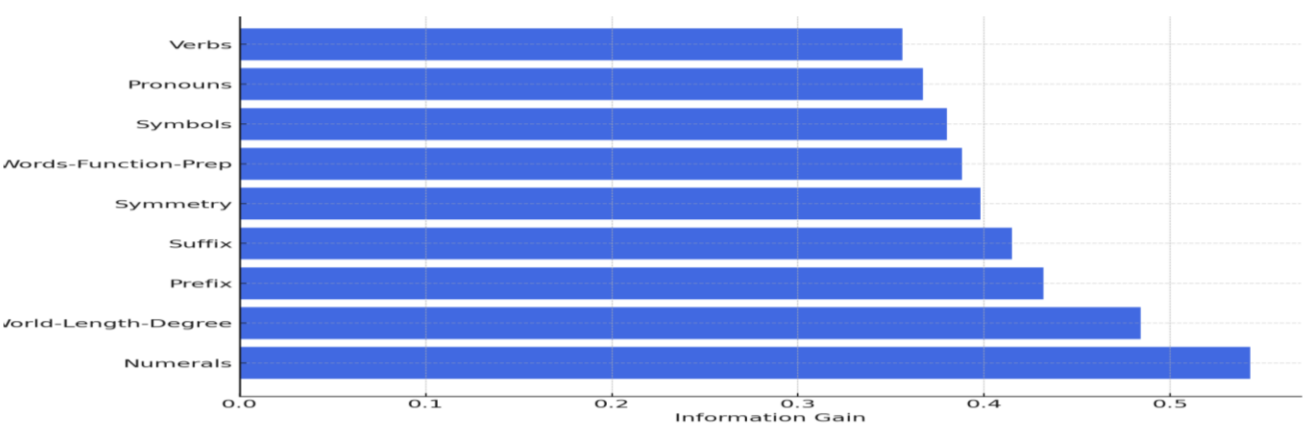

Fig. 7 Performance Metrics Across Queries

Fig. 7 represents the information gain of the twenty most discriminating linguistic features, such as "Numerals," "World-Length-Degree," and "Verbs."

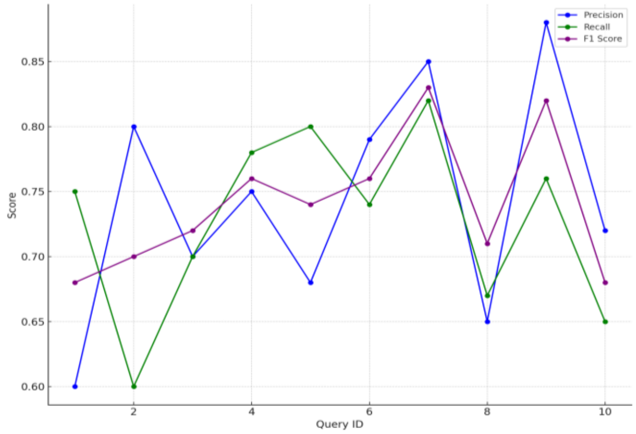

Fig. 8 Performance Metrics Across Queries

Fig. 8 shows a line chart comparing the performance metrics (Precision, Recall, F1 Score) across different queries.

IV. Conclusion

The shift from traditional keyword-based retrieval toward ontology-driven, semantically aware Filtering of information reflects a significant change in scholarly information systems. As discussed in the case study, the use of domain-specific ontologies enables richer representations of academic knowledge, which improves the precision, recall, and satisfaction levels among users. Disambiguation and contextual understanding of queries are possible with the incorporation of structured vocabularies and formal relationships among concepts, and as a result, highly relevant content is retrievable even from interdisciplinary fields. Advanced technologies such as OWL, RDF, and SPARQL allow for the processing of sophisticated semantic queries, dynamic classification of academic materials, and personalized retrieval. The comparative and case study analyses further demonstrate the significant increase in search accuracy, response time, and user interactions within Semantic Scholar, Europe PMC, and BioPortal. However, the introduction of ontology-based Filtering comes with challenges as well. In rapidly changing academic domains, designing, updating, and scaling ontologies requires sustained expert attention, making it difficult to maintain these systems. The use of legacy infrastructures alongside machine-deemed AI or automated machine-learning algorithms for ontology refinement presents essential challenges for these systems. Regardless of these challenges, the advantages of semantic filtering monitoring significantly outweigh the disadvantages, especially with the increased complexity of academic data. The next generation of scholarly information systems will need to employ hybrid semantic approaches that incorporate ontological reasoning together with adaptive systems in order to ensure relevance, scalability, and responsiveness. Such systems are increasingly required in academic communities for more automated and streamlined knowledge retrieval. As a result, ontology-driven approaches enable scholars to redefine interactions with voluminous scholarly resources that are rapidly evolving.

References

[1] Anand, J., Hemasundari, M., Kavitha Selvaranee, J., & Michael Mariadhas, J. (2024). Role of Strategic Human Resource Management and the Development of Information Systems for the Enhancement of Libraries. *Indian Journal of Information Sources and Services*, *14*(2), 78-84. https://doi.org/10.51983/ijiss-2024.14.2.12

[2] Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., ... & Sherlock, G. (2000). Gene ontology: tool for the unification of biology. *Nature genetics*, *25*(1), 25-29.

[3] Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., ... & Sherlock, G. (2000). Gene ontology: tool for the unification of biology. *Nature genetics*, *25*(1), 25-29. https://doi.org/10.1038/75556

[4] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), 993-1022.

[5] Musto, C., Semeraro, G., Lops, P., & De Gemmis, M. (2017). A System for Content-Based Recommendations. In F. Ricci, L. Rokach, & B. Shapira (Eds.), Recommender Systems Handbook (pp. 119–159). Springer.

[6] De los Ríos-Escalante, P. R., Contreras, A., Jara, P., Lara, G., Latsague, M., & Rudolph, E. (2023). A review knowledge of Chilean crayfishes (Decapoda: Parastacidae). *International Journal of Aquatic Research and Environmental Studies*, *3*(2), 109-115. https://doi.org/10.70102/IJARES/V3I2/7

[7] Doerr, M. (2003). The CIDOC conceptual reference module: an ontological approach to semantic interoperability of metadata. *AI magazine*, *24*(3), 75-75. https://doi.org/10.1609/aimag.v24i3.1720

[8] Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing?. *International journal of human-computer studies*, *43*(5-6), 907-928. https://doi.org/10.1006/ijhc.1995.1081

[9] Hearst, M. (2009). *Search user interfaces*. Cambridge university press.

[10] Horridge, M., Knublauch, H., Rector, A., Stevens, R., & Wroe, C. (2004). A practical guide to building OWL ontologies using the Protégé-OWL plugin and CO-ODE tools edition 1.0. *University of Manchester*.

[11] Musto, C., Gemmis, M. D., Lops, P., Narducci, F., & Semeraro, G. (2012). Semantics and content-based recommendations. In *Recommender systems handbook* (pp. 251-298). New York, NY: Springer US.

[12] Noy, N. F., & McGuinness, D. L. (2001). Ontology development 101: A guide to creating your first ontology. https://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html

[13] Noy, N., & McGuinness, D. L. (2001). Ontology development 101. *Knowledge Systems Laboratory, Stanford University*, *2001*, 1-18.

[14] Perera, T. D. P. (2018). Computer network analysis in knowledge sharing. *International Journal of Communication and Computer Technologies*, *6*(2), 5-8.

[15] Shamsfard, M., & Barforoush, A. A. (2003). The state of the art in ontology learning: a framework for comparison. *The Knowledge Engineering Review*, *18*(4), 293-316.

[16] Tanja, M., & Milica, V. (2023). The impact of public events on the use of space: Analysis of the manifestations in Liberty Square in Novi Sad. *Arhiv za tehničke nauke*, *2*(29), 75-82. https://doi.org/10.59456/afts.2023.1529.075M

[17] Tzitzikas, Y., Flouris, G., Papadakos, P. & Marketakis, Y. (2017). Ontology-based information retrieval: Survey and future directions. *Journal of Intelligent Information Systems, 52*(2), 371–409.