

An Intelligent Semantic-Aware Academic Library System Using Natural Language Processing and Knowledge Graphs

Mohamad Shady Alrahhah¹, Bandar Abdulwahed Fallatah², Omar F. Aloufi³, Abdullah M. Barakeh⁴
AIMD Laboratory-Computing College, Fahad Bin Sultan University, Tabuk City, KSA^{1,4}
The Boeing Company, Tabuk City, KSA²
Information Systems Department-Faculty of Computing and Information Technology,
King Abdulaziz University, Jeddah, Saudi Arabia³
Department of Information Systems-College of Computer Science and Engineering,
Taibah University, Madinah, Saudi Arabia³

Abstract—Traditional library systems designed for academic research suffer from poor integration when employing AI-based techniques for enhancement purposes. These limitations are mainly related to ignoring semantic relationships during search, overlooking hidden relationships among words, and lacking the generalization capability required to enhance researchers' experience. To address these issues, the Intelligent Library System (IntLS) is proposed and further enhanced through effective knowledge graph modeling to enable more accurate retrieval of results. In addition, the basic NLP processing steps are optimized to capture hidden relationships during tokenization, stop-word removal, and stemming stages. The architecture of the proposed system consists of two main components: a semantic component, responsible for generating semantic representations of documents, and an analytics component, responsible for analyzing historical searches to predict future user needs and support effective resource management. The proposed system is compared with two well-known systems using similarity-based accuracy and a set of AI-based evaluation metrics. The enhanced system, Enh-IntLS, demonstrates superior performance, achieving a 1.9% improvement in similarity-based accuracy and a 1.7% improvement in AI-based accuracy.

Keywords—Library systems; academic search; generalization; knowledge graph; similarity; accuracy

I. INTRODUCTION

Traditional Library Systems (TLSS) are essentially digital catalogues and circulation management platforms that provide limited analytics. They rely mainly on reactive usage statistics rather than predictive capabilities. In other words, traditional library management systems are database-driven systems that primarily digitize and automate manual processes. Such automation benefits stakeholders (i.e., both librarians and users) by enabling them to manage book records, track borrowing and returning transactions, monitor due dates, and generate basic reports [1]. In light of the rapid advancement of technologies, library system developers have begun employing artificial intelligence to enhance the quality of services provided to users. However, the integration of artificial intelligence also opens the door to new challenges.

A. Motivation

Despite these improvements, traditional library management systems still exhibit several limitations within modern information environments [2, 3]. First, their search mechanisms are primarily keyword-based, lacking semantic understanding and intelligent recommendation capabilities. As a result, users often struggle to find relevant resources beyond exact keyword matches. Second, these systems typically operate reactively rather than proactively; they do not predict user needs, recommend resources based on borrowing history, or dynamically optimize inventory. Third, traditional LMS platforms provide limited data analytics, which restricts libraries from extracting actionable insights related to user behavior, demand forecasting, and collection development. Furthermore, they lack automation in tasks such as intelligent cataloging, classification, anomaly detection (e.g., identifying overdue patterns), and user personalization. In many cases, decision-making remains dependent on manual librarian intervention, which reduces scalability and efficiency in large academic institutions.

B. Statement of Problem

With the advancement of Artificial Intelligence (AI), Machine Learning (ML), Natural Language Processing (NLP), and data analytics technologies, new opportunities have emerged to transform library services into intelligent, adaptive, and user-centric systems. AI-powered systems can enable semantic search, automated metadata generation, personalized recommendation engines, and predictive analytics for resource demand [4]. These capabilities can significantly enhance the user experience, improve operational efficiency, and support strategic resource management.

In other words, there remains a gap between conventional rule-based LMS platforms and fully intelligent, AI-driven library ecosystems. Many existing systems do not integrate advanced AI components within a unified and scalable architecture capable of supporting intelligent search, predictive decision-making, and adaptive service personalization. Bridging this gap is critical for establishing smart library systems that enable researchers to accurately find the documents they need. Moreover, to the best of our knowledge,

the evaluation of the generalization quality attribute has not been addressed in previous works, which is important for enhancing researchers' trust in academic search systems.

C. Research Questions

Based on the problem stated above, the following research questions are formulated:

- How can intelligent search be enabled based on the semantics of keywords?
- How can predictive analytics be enabled for resource demand?

D. Contributions

This work provides the following contributions:

- Proposes the Intelligent Library System (IntLS) to address the research questions mentioned above. The system employs an NLP-based approach in its development.
- Enhances IntLS by developing a knowledge graph of library topics, enabling accurate predictive analytics rather than relying solely on descriptive analytics.
- Evaluating the generalization quality attribute to demonstrate the effective applicability of the proposed system in real-world environments.
- Conducts a comparative evaluation with well-known smart library systems to demonstrate the superiority of the proposed approach.

E. Work Organization

The rest of this study is organized as follows: Section II presents an overview of both Traditional Library Systems (TLSs) and Intelligent Library Systems (ILSs). Section III describes the proposed system in detail. The experimental results are presented and discussed in Section IV. Finally, Section V concludes the study.

II. RELATED WORK

This section presents a brief overview of both traditional library systems and smart library systems.

A. Traditional Library Systems

Authors of work [5] stated that the need for a specialist academic library system is pressing to enhance the academic productivity of researchers. The previous work is supported by another article provided in [6], where a comparison between traditional and smart library systems is provided. The final result of this work stated that artificial intelligence techniques can play a significant role in enhancing the accuracy of academic search. The Ex Libris Alma platform [7], widely adopted by academic and research libraries worldwide, represents a recent generation of library services platforms that integrate cataloging, fulfillment, and resource management into a unified cloud-based system. According to the official Ex Libris documentation, Alma incorporates data analytics capabilities to improve operational efficiency in resource

management. However, its analytics remain primarily descriptive rather than predictive, and the platform relies on proprietary infrastructure with associated licensing costs that may not be accessible to all academic institutions, particularly those in developing regions.

B. Intelligent Library Systems

Many techniques are employed to develop smart library systems, such as text classification and natural language processing. In the work [8], the authors proposed a text classification-based system supported by the KNN algorithm. The key idea is to create clusters of text according to categories generated based on the smart recommendation of text classification. authors of work [9] presented an NLP-based library system supported by a correlation function to handle semantic issues. The key idea is to employ correlation based on the domain where the text belongs. The correlation function provides statistical information about semantics after following standard NLP steps. However, the limitation of this work is related to ignoring the knowledge graph database.

It is worth mentioning that KNN-based and NLP-based systems are involved in the comparison with the proposed system of this study.

III. PROPOSED SYSTEM

This section presents the proposed intelligent system, referred to as the Intelligent Library System (IntLS), as well as the enhanced one. First, the architecture of the system is introduced. Then, the role of each component within the architecture is described in detail. The testing of the system and the corresponding discussion of results are presented in a separate section.

A. IntLS Architecture

The architecture of IntLS is composed of two main components: the semantic component and the predictive analytics component. Fig. 1 below illustrates the role of the first component, in which several techniques are employed to generate the desired output.

To enable semantic-based search, the semantic component performs several steps, as described below.

1) *Semantic representation of documents*: The output of this step is to capture the meaning of documents rather than relying solely on exact word matches. This enables users to retrieve results based on context-aware queries. For example, if a user searches for books on artificial intelligence, resources containing topics such as deep learning, machine learning, and text mining will also appear in the result set. To achieve this capability, library resources (e.g., books, articles, theses, etc.) must first be transformed into semantic representations. To perform this task, text preprocessing is applied to important textual fields of the documents, such as the title, abstract, and keywords. Text preprocessing includes the following sub-tasks:

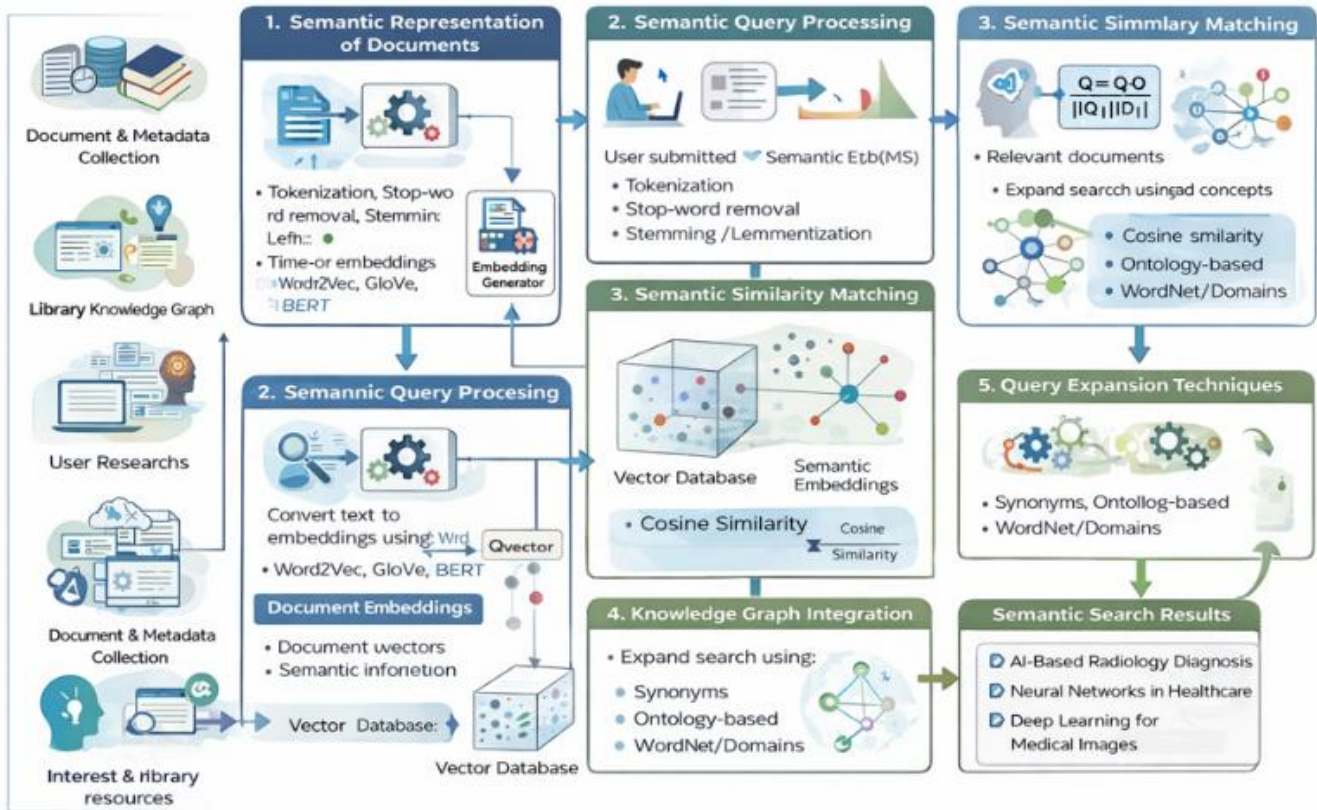


Fig. 1. Role of semantic component.

- **Tokenization:** This step involves converting a sentence into a sequence of words, commonly referred to as “tokens.” Tokens represent the basic building blocks upon which text analysis and other processing methods are performed. Using the title of work [10], the output will be as follows:

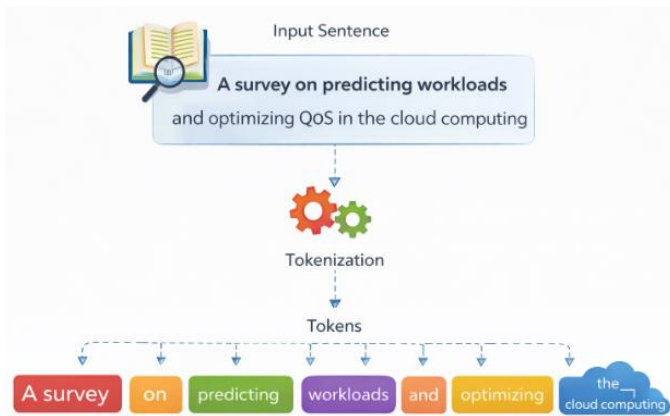


Fig. 2. Output of the tokenization step.

As shown in Fig. 2, a whitespace-based method is used to generate tokens. However, this work enhances this method by considering strong relationships between words rather than treating them individually. Therefore, the first token includes “A survey,” and the last token includes “cloud computing,” as they represent integral phrases within the computer science domain. In addition, abbreviations such as “QoS” in our

example are removed because they may introduce ambiguity and negatively affect accuracy.

- **Stop-word removal:** Researchers have provided a list of such words in the Bag-of-Words (BoW) library. The objective of this step is to eliminate words that have low importance in conveying meaning (i.e., words that have minimal impact on the overall semantics). Fig. 3 illustrates the output of the same previous example.

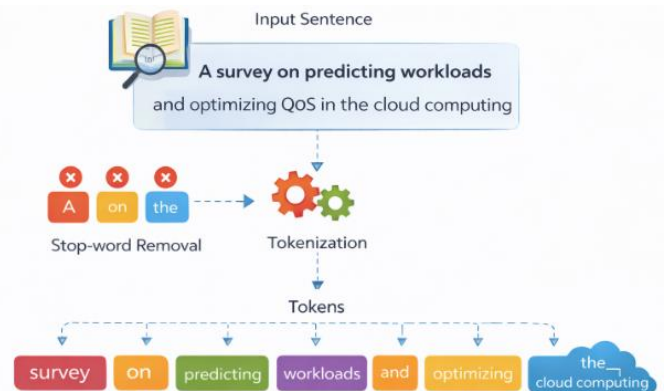


Fig. 3. Output of stop-word removal step.

As shown in Fig. 3, the words “A,” “on,” and “the” are removed. However, this work introduces an enhancement that considers the relationship conveyed by the word “and.” This is because authors of a document often use “and” to emphasize the relationship between two concepts (e.g., “workloads” and

“optimizing”), highlighting a form of trade-off or association between the two terms.

- **Stemming:** This is the process of converting words into their base form, or stem. In this study, Porter’s algorithm is used for this purpose. The output of this step is illustrated in Fig. 4.

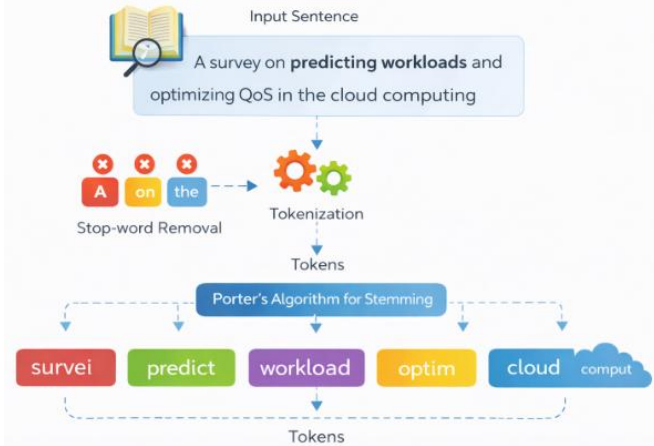


Fig. 4. Output of the stemming step.

- **Convert text into semantic vectors (embeddings):** This step uses the Word2Vec model to transform textual data into semantic vector representations, as shown in Fig. 5. The objective of this step is to map the text into a numerical vector space, enabling similarity calculations in subsequent stages.

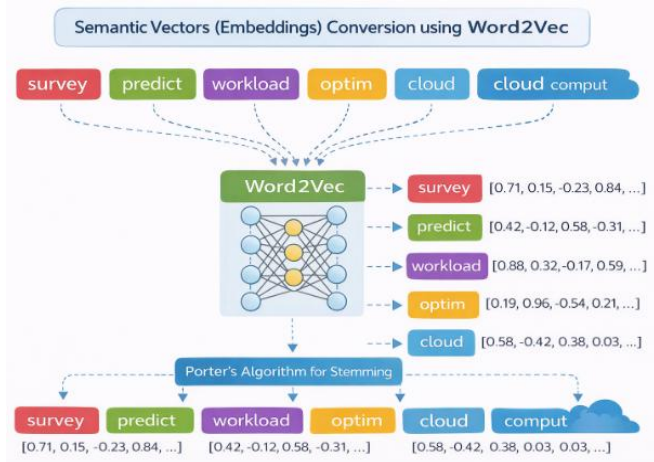


Fig. 5. Output of embedding step.

2) *Semantic query processing:* The objective of this step is to enable IntLS to retrieve documents with similar meanings even when different keywords are used. For example, if a user submits the query “intelligent learning for the medical field,” the system can retrieve studies such as “deep learning in medical imaging” or “neural network–based COVID-19 diagnosis.”

In this context, the same preprocessing steps used in the previous example are applied (i.e., tokenization, stop-word removal, stemming, etc.). The query is then converted into a

semantic embedding vector. Finally, semantic similarity matching is performed, as described below.

3) *Semantic similarity matching:* The objective of this step is to rank the retrieved documents within the result set based on their semantic closeness to the query. In this context, the Cosine Product similarity measure [11] is used. It is defined as follows:

$$Sim(D, Q) = \frac{D \cdot Q}{\|D\| \times \|Q\|} \quad (1)$$

where, Q is the query vector, and D is the document vector.

Problem. A critical issue arises when considering the intersections among different domains. For example, the computer vision domain is tightly coupled with deep learning, while artificial intelligence is strongly linked with text and data mining. In such scenarios, the accuracy of retrieval results may decrease due to overlapping or related concepts across domains. To address this challenge, this work introduces an enhancement to IntLS through the integration of a Knowledge Graph [12].

B. Enhanced IntLS Architecture (Enh-IntLS)

The key idea behind this enhancement is the construction of a knowledge graph of library topics. The objective of this enhancement is to support semantic search, recommendation, topic discovery, and trend analysis across intersecting domains, which is essential for researchers seeking relevant documents. A knowledge graph is a network of nodes and edges, where nodes represent entities such as authors, topics, journals, and departments, while edges represent relationships such as *written by*, *belongs to*, *related to*, and *published in*. Through this structure, Enh-IntLS can understand how topics, books, authors, users, and research areas are interconnected. Fig. 6 illustrates the basic steps involved in creating the knowledge graph.

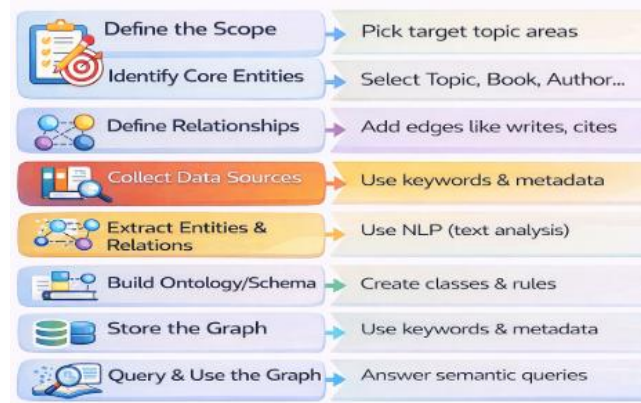


Fig. 6. Steps of knowledge graph development.

From mathematical modeling, the basic steps are provided as follows:

$$\text{Let the final knowledge graph be: } G = (V, E) \quad (2)$$

where, V is the set of entities (nodes), and E is the set of relationships (edges). The scope $S \in D$, where D is all possible domains, such as cloud computing.

The graph schema includes the following entity types:

$V = \{\text{Publication, Author, Topic, FieldOfStudy, Journal, Conference, Department, Institution}\}$

The main relation types are defined as follows:

$E = \{\text{written_by, published_in, belongs_to, has_topic, cites, affiliated_with, related_to}\}$

Important entities extracted from documents are represented by:

$$E = \{e_1, e_2, e_3, \dots, e_n\} \quad (3)$$

Entities are extracted from structured and semi-structured academic metadata. The main input fields include publication titles, abstracts, keywords, authors, publication venues, departments, institutions, citation links, and fields of study. The extraction process follows these steps: First, the textual fields of each document are preprocessed using tokenization, stop-word removal, phrase detection, and stemming. Second, candidate entities are identified from titles, abstracts, and keywords. Third, structured metadata fields, such as author names, journal names, conference names, institutional affiliations, and fields of study, are directly extracted as graph entities.

After extraction, each entity is linked to an existing graph node or inserted as a new node if no match is found. Entity linking is performed using three steps. First, entity labels are normalized by converting them to lowercase, removing punctuation, and standardizing abbreviations. Second, exact matching is used to link identical entity names. Third, semantic matching is applied using embedding similarity to link conceptually equivalent or highly related entities.

Between two different entities, the relationship is defined as:

$$E = \{(e_i, r, e_j)\} \quad (4)$$

where, e_i, e_j are entities, and r is kind of relationship.

To collect data from different resources, let:

$$X = \{x_1, x_2, x_3, \dots, x_k\} \quad (5)$$

Refers to the set of resources. Then, each document x can be represented by:

$$x = \{\text{title, abstract, keywords}\} \quad (6)$$

E_i is the extracted entity from a document x . Then, the relationship extracted from the pairs is:

$$r = \text{relation}(e_i, e_j) \quad (7)$$

Relations are extracted using both metadata-based and semantic-based methods. Metadata-based relations are obtained directly from structured fields. For example, if publication (P_i) has author (a_i) the triple ($P_i, \text{written_by}, a_i$) is created.

The relationships are embedded within the graph. However, the graph is stored using three parameters as follows:

$$\text{StrG} = (s, p, o) \quad (8)$$

where, S is the subject, P is the predicate, and O is the object.

For example, (Article A covers Cloud Computing). Based on this representation, queries are issued from the side of the user, so that searching relevant nodes for a query is modeled as:

$$\text{Result} = \text{argmax}_{e \in E} \text{Score}(q, e) \quad (9)$$

To make the enhancement clearer, the following provides an example of query enrichment: Assume that a user submits the query: “deep learning in medical imaging”. First, the query is preprocessed and converted into a semantic vector. The embedding component retrieves candidate documents that contain semantically close expressions such as “neural networks for diagnosis,” “CNN-based image classification,” and “AI-based radiology analysis.” Then, the query entities, deep learning and medical imaging, are linked to corresponding nodes in the knowledge graph. Through graph traversal, Enh-IntLS identifies related nodes such as neural networks, image classification, medical diagnosis, radiology, and computer vision. Candidate documents connected to these related nodes receive higher graph-based relevance scores. For example, a document titled “Neural Network-Based COVID-19 Diagnosis from Chest Images” may be ranked higher even if it does not contain the exact phrase “deep learning in medical imaging.”

After the enhancement based on the knowledge graph, the predictive analytics component begins its function. The main objective of this component is to analyze historical usage data and semantic information about library resources to forecast future trends and user needs. To accomplish this task, the predictive analytics component performs the steps described below.

1) *Predicting user information needs:* By analyzing previous searches and browsing history, a prediction profile is established to identify resources that users may need in the future.

2) *Demand forecasting for library resources:* This step predicts which books, journals, or digital resources will be highly in demand in the future.

Mathematically, these actions are modeled as follows:

Let UP is user profile features, RF is resource features (topics, keywords, metadata), HID is historical interaction data, and f is predictive model. Then, predicted demand (PD) is:

$$PD = f(UP, RF, HID) \quad (10)$$

Fig. 7 illustrates the integration of the two components to generate the final output.



Fig. 7. Integration of the semantic component and predictive analytics component.

IV. EXPERIMENTAL RESULTS AND EVALUATION

To evaluate the proposed IntLS and Enh-IntLS systems, the most recent Microsoft Academic Graph (MAG) dataset [13] was selected for training and testing. The dataset contains 109,105 triples. It is a heterogeneous graph that includes scientific publication records, citation relationships among these publications, as well as information about authors, institutions, journals, conferences, and fields of study. Therefore, it is considered a suitable benchmark dataset for conducting experiments. The dataset was divided into 80% for training and 20% for testing.

Evaluation is arranged to cover two main tasks, which are: 1) Semantic retrieval of academic resources; and 2) Resource demand prediction. For each task, input data, target output, ground truth labels, and evaluation metrics are defined as follows:

1) For the semantic retrieval task, the input consists of academic resource metadata, including titles, abstracts, keywords, topics, and knowledge-graph triples extracted from the Microsoft Academic Graph dataset. The target output is a ranked list of academic resources relevant to a submitted query. The relevance criterion is based on semantic overlap between the query concepts and the document concepts represented through Word2Vec embeddings and knowledge-graph relationships. The evaluation uses similarity-based accuracy to measure whether the retrieved resources are semantically aligned with the expected relevant topics.

2) For the demand prediction task, the input consists of simulated historical interaction records derived from academic-resource usage indicators, including resource-topic frequency, query-topic occurrence, and resource-demand class labels. The target output is a binary classification of future demand, namely correct/high demand and incorrect/low demand. The ground truth labels are represented through demand classes used in the confusion matrix. The evaluation uses accuracy, precision, recall, specificity, and ROC-based analysis.

It is worth mentioning that since the Microsoft Academic Graph is primarily a scholarly metadata and citation graph, it is integrated with simulated demand data to cover the evaluation of the second task. In other words, simulated interaction and synthetic demand-label data are used for predictive analytics. Specifically, simulated historical interaction records were constructed from MAG-based academic-resource indicators. These indicators included topic frequency, resource-topic occurrence, citation activity, field-of-study frequency, and graph-based relationships among publications, authors, and research areas.

This integration leads to update number of records in the selected dataset to be 214,000 records. In other words, the value of 214,000 evaluated instances refers to simulated demand-observation records generated from resource-topic occurrences, citation activity, field-of-study frequency, and synthetic interaction indicators. These simulated records were used only for evaluating the demand-prediction task, not for evaluating the semantic retrieval task.

B. Strategy of Evaluation and Used Metrics

We first evaluate the main system (i.e., IntLS) in comparison with the enhanced system (i.e., Enh-IntLS). Then, Enh-IntLS is compared with KNN-based and NLP-based systems reviewed in the related work section. In addition, the generalization quality attribute is evaluated by varying the training-to-testing ratios. It is worth mentioning that in this work, generalization is considered internal validation, as no real data is available. In future work, we tend to prove generalization using real data. For forecasting resource demand, the logistic regression algorithm [14] is employed.

Regarding the evaluation metrics, similarity-based accuracy is used to assess semantic similarity matching based on Eq. (1). Specifically, we define the semantic retrieval task as a top-k ranking problem. For each query, the system retrieves and ranks academic documents according to the semantic similarity score between the query vector and the document vector. In IntLS, this score is computed using Word2Vec-based document/query embeddings and Cosine Product similarity. In Enh-IntLS, the similarity score is further enriched using knowledge-graph relationships among topics, authors, journals, departments, and publications. A retrieved document is considered relevant if its semantic similarity score is greater than or equal to a predefined threshold and if it shares at least one expected topic, field-of-study, or knowledge-graph relationship with the query. In the revised manuscript, we define the threshold value as **0.70** after normalizing similarity scores into the range [0, 1]. Therefore, a retrieved document is considered correct when it satisfies the relevance criterion, and it is considered incorrect when it does not satisfy this criterion.

In addition, accuracy, precision, recall, and specificity are used to evaluate the prediction of resource demand. These metrics are derived from the confusion matrix [15] summarized in Table I.

TABLE I. CONFUSION MATRIX

Predicted Dem Actual Dem	PD _{final}	\neg PD _{final}	Sum
PD _{final}	TP	FN	S
\neg PD _{final}	FP	TN	Su
Sum	\bar{S}	\bar{S}_u	All

where:

- True positives (TPs): The positive demands that were correctly recommended by the system.
- True negatives (TNs): The negative demands that were correctly recommended by the system.
- False positives (FPs): The negative demands that were incorrectly recommended as positives by the system.
- False negatives (FNs): The positive demands that were recommended as negatives by the system.

$$\text{Accuracy} = \frac{TP+TN}{All} \quad (11)$$

$$\text{Precision} = \frac{TP}{\bar{S}} \quad (12)$$

$$\text{Recall} = \frac{TP}{S} \quad (13)$$

$$\text{Specificity} = \frac{TN}{S_u} \quad (14)$$

C. Similarity-Based Accuracy Evaluation of IntLS and Enh-IntLS (Semantic Retrieval Task)

In this subsection, the semantic retrieval task is evaluated, and the corresponding results are discussed. Table II presents the results obtained under varying training-to-testing ratios for evaluating the generalization quality attribute.

TABLE II. SIMILARITY-BASED ACCURACY RESULTS

System	Training to Testing Ratio	Sim Metric	AVG
IntLS	[80 to 20]	80.2	≈ 78.58
	[70 to 30]	80.1	
	[60 to 40]	79	
	[50 to 50]	75	
Enh-IntLS	[80 to 20]	93.3	≈ 88.83
	[70 to 30]	91	
	[60 to 40]	88	
	[50 to 50]	83	

Discussion. When the training ratio was decreased by 10%, and the testing ratio was increased by 10%, both systems showed decreasing values in the similarity-based accuracy metric. This can be explained by the fact that incorporating more records into the training dataset leads to the extraction of more features, which can be utilized during the similarity matching process, thereby resulting in higher accuracy values. However, Enh-IntLS demonstrated better performance compared to IntLS. This improvement can be attributed to the use of a knowledge graph of topics, which models similarity based on the semantics of input queries and retrieved results. Regarding the generalization quality attribute, Fig. 8 illustrates the difference between the baseline training-to-testing ratio and the average results obtained for both systems.

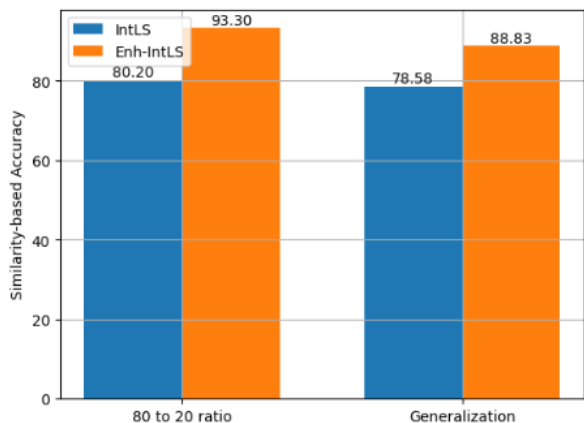


Fig. 8. Results of similarity-based accuracy metric.

As shown in Fig. 8, the difference between the baseline training-to-testing ratio and the generalization results is small (approximately 1.4% for IntLS and about 4% for Enh-IntLS).

This indicates that the systems demonstrate acceptable generalization capability and can be applied in real-world environments to enhance researchers' trust in the system.

D. Accuracy, Precision, Recall, and Sensitivity Evaluation of IntLS and Enh-IntLS (Demand Prediction Task)

In this subsection, the demand prediction task is evaluated, and corresponding results are discussed. Fig. 9 and 10 illustrate confusion matrix visualizations for both systems.

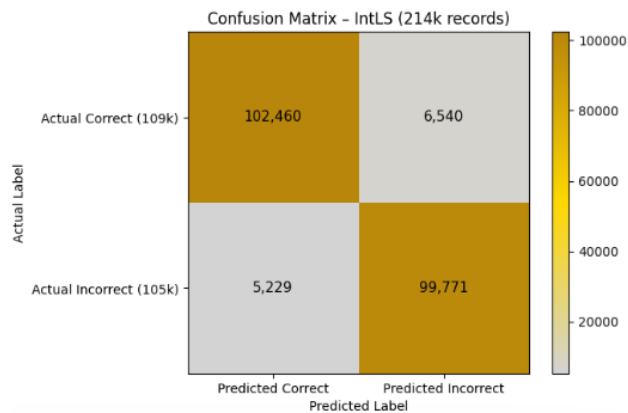


Fig. 9. Confusion matrix values of IntLS.

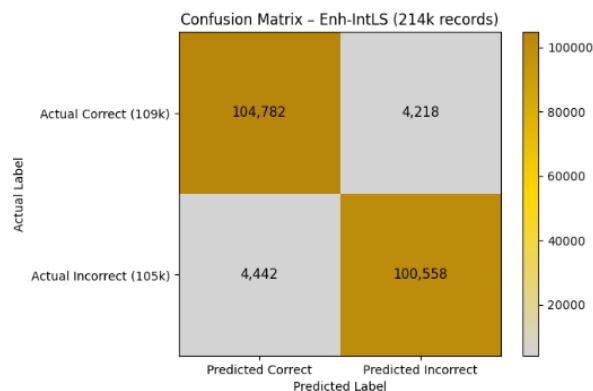


Fig. 10. Confusion matrix values of Enh-f IntLS.

Table III summarizes the obtained results without involving the generalization quality attribute.

TABLE III. RESULTS OF AI-BASED METRICS [80 TO 20] RATIO

System	Metric	AVG Value	Class	
			Correct Demand	Incorrect Demand
IntLS	Accuracy	94.50	92.34	96.66
	Precision	95.14	93.4	96.88
	Recall	94.00	94	94.00
	Specificity	95.02	95.02	95.02
Enh-IntLS	Accuracy	95.95	96.22	95.68
	Precision	95.93	95.89	95.97
	Recall	96.13	96.13	96.13
	Specificity	95.77	95.77	95.77

As shown in Table III, the value of each metric is presented in detail according to the correct and incorrect demand classes. It is evident that Enh-IntLS achieves better scores compared to IntLS. This improvement can be attributed to the fact that resource demand encompasses a wide spectrum of semantics that are not captured by IntLS. Since accuracy is considered one of the most important metrics in the AI domain, generalization is illustrated in Fig. 11.

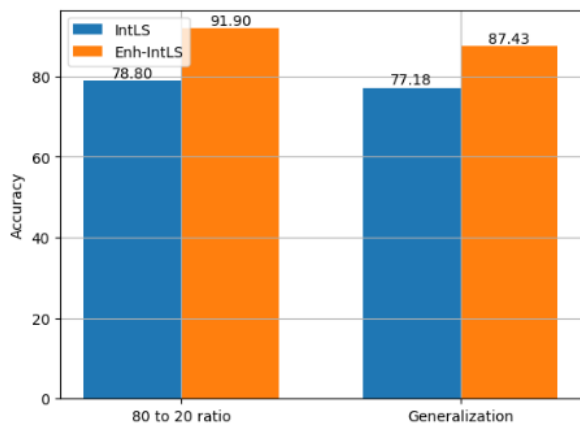


Fig. 11. Results of the accuracy metric.

As shown in Fig. 11, the difference between the baseline training-to-testing ratio and the generalization results is again small, which provides further evidence of the generalization capability of the proposed systems. However, Enh-IntLS records better values for the same reason discussed earlier when presenting the results of the similarity-based accuracy metric.

E. Results of Comparison to KNN-Based and NLP-Based Systems

In this context, only Enh-IntLS is involved in the comparison process, as it achieves better results than IntLS. For the similarity-based accuracy metric and under the generalization criterion, the previously obtained results of Enh-IntLS are considered for comparison with the selected systems. Fig. 12 illustrates the comparison results, where the KNN and NLP systems are implemented to generate a simplified version using the same selected dataset.

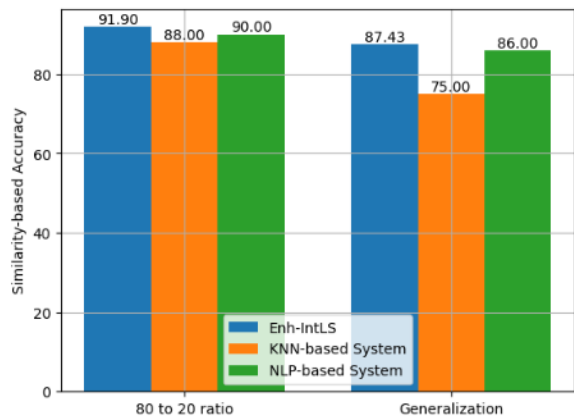


Fig. 12. Similarity-based accuracy metric comparison.

As shown in Fig. 12, the KNN-based system achieves the lowest performance compared to the other systems, as it does not incorporate knowledge graphs within its structure. This limitation leads to the retrieval of irrelevant documents within the result set. The NLP-based system achieves values close to those of the proposed system, as it employs similar NLP processing stages to those used in the development of Enh-IntLS. However, Enh-IntLS still ranks at the top because it incorporates more effective mathematical modeling within its architecture. In addition, the relationships conveyed through linking words, such as “and,” are not captured by the NLP-based system. Numerically, the Enh-IntLS achieves 1.9-based similarity-based accuracy enhancement.

Regarding the accuracy metric, which is one of the most important metrics in the AI domain, the ROC curve illustrated in Fig. 13 presents the obtained results. The ROC curve reflects the classification performance by comparing different classifiers based on the area under the curve (AUC) [16].

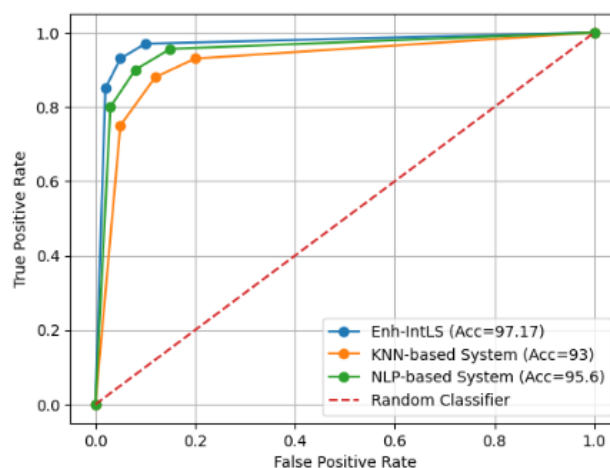


Fig. 13. ROC curve based evaluation.

As shown in Fig. 13, the Enh-IntLS system achieves the best performance, with its ROC curve located closest to the top-left corner, indicating a higher true positive rate and a lower false positive rate. The NLP-based system also demonstrates strong performance but remains slightly below the Enh-IntLS curve. In contrast, the KNN-based system shows comparatively lower performance, as its curve is farther from the optimal region. Overall, the ROC analysis confirms that Enh-IntLS provides superior discriminative capability and more reliable prediction accuracy compared to the other evaluated approaches.

V. CONCLUSION

This study proposed an Intelligent Library System (IntLS) designed to improve academic search and resource management by incorporating semantic understanding and predictive analytics. The system addresses key limitations of traditional library management systems, particularly the lack of semantic awareness, limited exploitation of relationships among research topics, and insufficient support for predictive decision-making. To overcome these challenges, an enhanced version of the system, Enh-IntLS, was developed by integrating knowledge graph modeling with optimized NLP

preprocessing techniques to capture hidden relationships among words and research topics. The architecture of the proposed framework combines a semantic component for generating meaningful document representations and a predictive analytics component for forecasting future resource demand based on historical usage patterns. The evaluation strategy is structured such that IntLS and Enh-IntLS are assessed using similarity-based accuracy, precision, recall, and sensitivity metrics. Subsequently, the best results obtained by Enh-IntLS are compared with two similar existing approaches to demonstrate its superiority. Experimental evaluation using the Microsoft Academic Graph dataset demonstrated that Enh-IntLS outperforms both the baseline IntLS and other comparison systems, achieving improvements in similarity-based accuracy and AI-based evaluation metrics. Furthermore, the results confirmed that the proposed systems maintain acceptable generalization capability under varying training-to-testing ratios. These findings indicate that the proposed approach can significantly enhance the effectiveness of academic library search systems and support better resource discovery and management for researchers.

Some of the non-functional requirements, as illustrative screenshots, are presented in the Appendix.

VI. LIMITATIONS AND FUTURE WORK

The main limitation of this study is that computational complexity was not analyzed, and cross-validation techniques were not employed to evaluate their impact on accuracy. Furthermore, only one dataset was used to train and test the proposed system. In future work, these limitations will be addressed by providing a detailed analysis of computational complexity and by using multiple datasets for training and testing to further validate the robustness of the proposed approach. In addition, generalization is considered as internal validation, and in future work, we tend to prove generalization using real data.

ACKNOWLEDGMENT

The authors used ChatGPT for proofreading and generating illustrative figures. All scientific content, analysis, and conclusions were developed by the authors.

REFERENCES

- [1] Bagewadi, Pooja, et al. "Automating Libraries: The role of library management systems in modern library operations." *International Journal of Research in Library Science* 11.2 (2025): 105-114.
- [2] Liu, Ze, and Bo Shao. "A systematic review of library services platforms research and research agenda." *Library & Information Science Research* 46.4 (2024): 101325.
- [3] Khalid, Ayesha, Ghulam Farid Malik, and Khalid Mahmood. "Sustainable development challenges in libraries: A systematic literature review (2000–2020)." *The Journal of academic librarianship* 47.3 (2021): 102347.
- [4] Kanumarlapudi, Praveen Kumar. "AI-Powered Product Metadata Enrichment through a Hybrid Approach Combining Semantic Web and Machine Learning." *Learning* 2.2 (2025): 1-17.
- [5] Cox, Andrew M., and Sheila Corral. "Evolving academic library specialties." *Journal of the American Society for Information Science and Technology* 64.8 (2013): 1526-1542.
- [6] Yudov, Andriy. "Traditional and modern library services: Transformation." *Library Science. Record Studies. Infomology* 3.19 (2023): 34-41.

- [7] Ex Libris Web Site. Available [Online]: <https://exlibrisgroup.com/products/alma-library-services-platform/>
- [8] Wang, Changjun, Fengxia You, and Yu Wang. "Text intelligent classification model based on improved KNN algorithm in library service transformation." *Systems and Soft Computing* 7 (2025): 200313.
- [9] Sawant, Pradnya, and Kavita Sonawane. "NLP-based smart decision making for business and academics." *Natural Language Processing Journal* 8 (2024): 100090.
- [10] Aloufi, Omar F., et al. "A survey on predicting workloads and optimizing QoS in the cloud computing." 2021 International Congress of Advanced Technology and Engineering (ICOTEN). IEEE, 2021.
- [11] Naveen, L., Ananda Raj, and S. Rajalakshmi. "Abstractive text summarizer: A comparative study on dot product attention and cosine similarity." 2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT). IEEE, 2021.
- [12] Xiong, Chenyan, Russell Power, and Jamie Callan. "Explicit semantic ranking for academic search via knowledge graph embedding." *Proceedings of the 26th international conference on world wide web*. 2017.
- [13] Microsoft Academic Graph dataset Website. Available [Online]: <https://www.microsoft.com/en-us/research/project/microsoft-academic-graph> (Accessed: 16/03/2026)
- [14] Vimal, Bhartendoo, and S. Anupama Kumar. "Application of logistic regression in natural language processing." *Int J Eng Res* 9.06 (2020).
- [15] Hsu, Li-Yuan, et al. "Toward building an academic search engine understanding the purposes of the matched sentences in an abstract." *IEEE Access* 9 (2021): 109344-109354.
- [16] Janssens, A. Cecile JW, and Forike K. Martens. "Reflection on modern methods: Revisiting the area under the ROC Curve." *International journal of epidemiology* 49.4 (2020): 1397-1403.

APPENDIX

This appendix is provided to cover some of the non-functional requirements as illustrative screenshots.

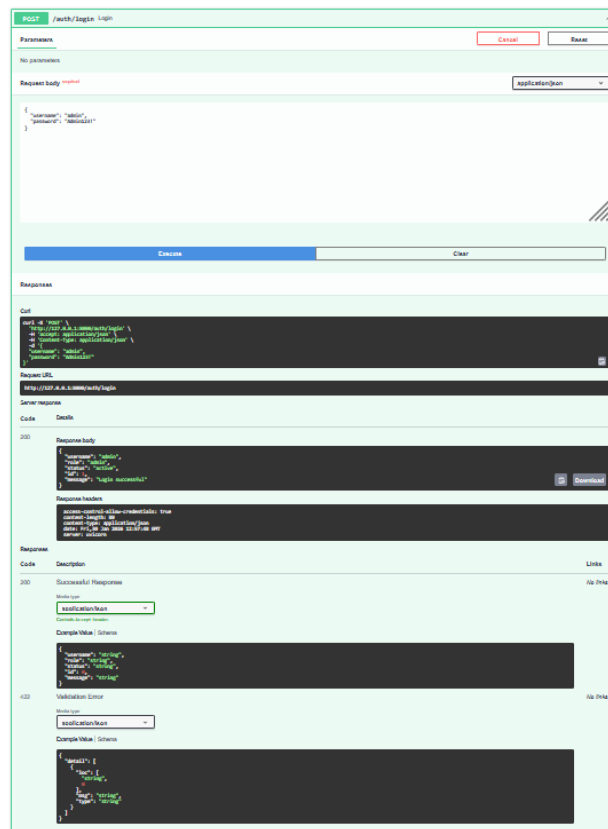


Fig. 14. Login interface.

User authentication successfully validated credentials against securely hashed password storage using bcrypt. Session management maintains user state across multiple requests without requiring repeated authentication. Role-based access control correctly restricted administrative operations to authorized users while permitting standard operations for all authenticated users. Fig. 14 demonstrates successful user authentication through the system login interface.

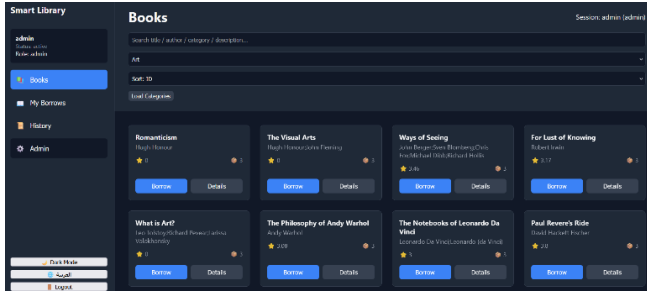


Fig. 15. Browsing interface.

The catalog search functionality correctly retrieved bibliographic records matching user-specified queries, filters, and sorting criteria. Full-text search operated across the title, author, category, and description fields. Filtering by category and sorting by relevance, title, or publication year produced the expected result sets. Fig. 15 illustrates the catalog browsing interface with active search and filter controls.

Borrowing operations correctly decremented the available copy counts and created borrowing records with appropriate due dates. The system prevented borrowing when no copies were available and enforced borrowing limits per user. Return operations restored a availability, updated the borrowing status, and calculated overdue periods when applicable, as shown in Fig. 16.

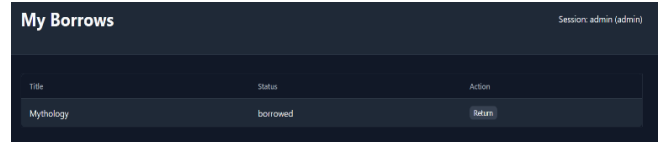


Fig. 16. Update borrowing interface.