

ARTIFICIAL INTELLIGENCE, LARGE LANGUAGE MODELS, AND SEMANTIC WEB TECHNOLOGIES FOR NEXT-GENERATION DIGITAL LIBRARIES

(A Hybrid Framework for Intelligent Knowledge Discovery)

Dr. Karunakar N.

College Librarian (SG)
Government College for Women (Autonomous), Mandya, Karnataka, India

Dr. Nagesh R.

College Librarian
Government First Grade College, Bangarpet, Kolar District, Bengaluru District, Karnataka, India

ABSTRACT

Digital libraries are rapidly evolving from conventional repositories into intelligent knowledge environments capable of providing contextual, personalized, and semantically enriched information services. The integration of Artificial Intelligence (AI), Large Language Models (LLMs), Semantic Web technologies, and Knowledge Graphs offers significant opportunities to improve information retrieval and intelligent knowledge discovery. This study proposes a hybrid framework that integrates AI and LLM-based natural-language processing with semantic knowledge representation, ontologies, metadata, and Knowledge Graphs. The proposed framework combines conventional keyword retrieval, semantic search, and AI-based contextual processing to improve the accuracy and relevance of retrieved information. It further incorporates intelligent question answering, summarization, recommendation, and personalized user interaction. The methodology adopts a systematic and analytical research design using scholarly literature, digital library resources, metadata, and suitable datasets. The framework is evaluated using retrieval accuracy, relevance, usability, user satisfaction, and knowledge-discovery measures, including precision, recall, and F1-score. The study also addresses major challenges associated with AI-enabled digital libraries, including hallucination, reliability, algorithmic bias, privacy, data security, copyright, explainability, and human oversight. The proposed framework contributes to Library and Information Science by providing an integrated technological model for next-generation digital libraries. It demonstrates how the convergence of AI, LLMs, and Semantic Web technologies can support more intelligent, context-aware, personalized, and knowledge-oriented information discovery while emphasizing responsible and trustworthy AI adoption.

Keywords: Artificial Intelligence, Large Language Models, Digital Libraries, Semantic Web, Knowledge Graphs, Intelligent Information Retrieval, Knowledge Discovery, Generative AI

1. INTRODUCTION

1.1 Background of the Study

Digital libraries are evolving from conventional repositories of digitized information into intelligent knowledge environments capable of supporting advanced discovery and personalized information access. The integration of Artificial Intelligence (AI), Large Language Models (LLMs), Semantic Web technologies, and Knowledge Graphs provides new opportunities to transform information retrieval and knowledge discovery. AI and LLMs

can support natural-language interaction, contextual understanding, content summarization, and intelligent recommendations, while Semantic Web technologies enable meaningful relationships among heterogeneous information resources. The proposed study therefore focuses on developing a hybrid framework that combines these technologies to create next-generation digital libraries with improved retrieval accuracy, relevance, personalization, and intelligent knowledge discovery.

1.2 Problem Statement

Despite significant developments in digital library technologies, conventional information retrieval systems often depend heavily on keyword-based searching and may have limitations in understanding user intent, contextual relationships, and the semantic meaning of information resources. Existing AI-enabled systems may provide intelligent responses, but concerns regarding hallucination, reliability, bias, privacy, copyright, and explainability remain important. Similarly, isolated implementation of AI, LLMs, or Semantic Web technologies may not fully address the complex requirements of modern knowledge discovery. The central problem of this study is therefore the absence of an integrated and systematically evaluated framework that combines AI, LLMs, Semantic Web, and Knowledge Graph technologies for intelligent digital library retrieval and discovery.

1.3 Research Gap

The existing literature demonstrates growing applications of AI in digital libraries, increasing use of Large Language Models and Generative AI, and the development of Semantic Web and Knowledge Graph approaches for intelligent information retrieval. However, these technological approaches are frequently examined separately rather than through a unified hybrid framework. There remains a need to systematically integrate AI and LLM capabilities with semantic knowledge representation and knowledge graphs to improve contextual retrieval, relevance, personalization, and knowledge discovery. Furthermore, comprehensive evaluation comparing such a hybrid approach with conventional retrieval systems remains limited. This study addresses this gap by proposing and evaluating a hybrid AI–LLM–Semantic Web framework for next-generation digital libraries.

1.4 Research Objectives

1. To examine the role and applications of Artificial Intelligence (AI), Large Language Models (LLMs), Semantic Web technologies, and Knowledge Graphs in next-generation digital libraries.
2. To develop a hybrid framework integrating AI, LLMs, Semantic Web, and Knowledge Graph technologies for intelligent information retrieval and knowledge discovery.
3. To evaluate the proposed hybrid framework in terms of retrieval accuracy, relevance, user satisfaction, usability, personalization, and intelligent knowledge discovery in comparison with conventional retrieval approaches.

1.5 Research Questions

1. What are the roles and applications of AI, LLMs, Semantic Web technologies, and Knowledge Graphs in next-generation digital libraries?
2. How can AI, LLMs, Semantic Web technologies, and Knowledge Graphs be integrated into a hybrid framework for intelligent information retrieval and knowledge discovery?

3. How effective is the proposed hybrid framework compared with conventional retrieval systems in terms of accuracy, relevance, usability, user satisfaction, and knowledge discovery?

1.6 Significance of the Study

The study is significant because it proposes an integrated approach to transforming digital libraries into intelligent knowledge discovery environments. By combining Artificial Intelligence, Large Language Models, Semantic Web technologies, and Knowledge Graphs, the study can enhance information retrieval, contextual understanding, relevance, personalization, and user interaction. The proposed framework may help library professionals and digital library developers design more effective intelligent information services. It also contributes to Library and Information Science research by addressing the integration of emerging technologies within digital library systems. Furthermore, the study highlights important issues related to AI reliability, algorithmic bias, privacy, copyright, explainability, and human oversight, supporting responsible technological adoption.

2. REVIEW OF LITERATURE

2.1 AI in Digital Libraries

Das and Islam (2021) conducted a systematic review of AI and Machine Learning applications in libraries. Their study reviewed 32 selected articles from major scholarly databases and found that AI research in Library and Information Science was largely focused on theoretical developments, although implementation projects and case studies were also emerging. The authors highlighted the potential of AI to improve library services, information retrieval, and technology-oriented practices. Their findings demonstrate the increasing importance of AI for developing intelligent library systems and provide a foundation for integrating AI technologies into next-generation digital libraries.

Martínez Concha, Palacios Zenteno, and Tello Alfaro (2024) conducted a systematic review of research on AI applications in libraries during 2019–2023. Their study examined the changing characteristics of AI research and its increasing relevance to library services. The review indicates that AI is becoming an important technological component in modern library environments, with applications aimed at improving services and supporting information-related activities. The study also demonstrates the continuing development of AI research in libraries and the need to examine emerging applications systematically as libraries transition toward more technologically advanced and intelligent information environments.

2.2 Large Language Models and Generative AI

Kumar (2024) examined the development, architecture, applications, and challenges of Large Language Models (LLMs). The study highlighted the capabilities of models such as GPT-4, PaLM, and other advanced language models in natural language understanding, generation, reasoning, and knowledge-oriented applications. Kumar emphasized that the rapid development of LLMs has created significant opportunities across different domains while simultaneously raising concerns regarding reliability, computational requirements, bias, and ethical use. The findings are relevant to digital libraries because LLMs can facilitate natural-language interaction, intelligent information access, summarization, and knowledge discovery.

Zhu et al. (2023) reviewed the integration of Large Language Models with Information Retrieval (IR) systems and examined their applications in query rewriting, retrieval, reranking, and reading. The authors observed that LLMs provide strong contextual and semantic understanding, enabling information retrieval systems to move beyond traditional

keyword-based approaches. However, challenges related to interpretability, data requirements, computational costs, and the generation of plausible but inaccurate information remain. The study suggests combining conventional retrieval methods with LLM-based approaches to improve information access and reliability. This perspective directly supports the present study's focus on integrating LLMs with intelligent retrieval and digital library knowledge discovery.

2.3 Semantic Web and Knowledge Graphs

Hogan et al. (2021) presented a comprehensive study of Knowledge Graphs, examining their role in representing, integrating, querying, and extracting knowledge from large and heterogeneous datasets. The authors discussed graph-based data models, query and validation languages, ontologies, and techniques for knowledge representation and extraction. Their study highlighted the importance of Knowledge Graphs for connecting entities and relationships in a structured manner, thereby supporting advanced information retrieval and intelligent knowledge discovery. The findings are particularly relevant to digital libraries, where heterogeneous information resources can be semantically connected to improve contextual searching and retrieval.

Ji et al. (2021) conducted a comprehensive survey of Knowledge Graph representation, acquisition, and applications. The study examined knowledge graph representation learning, knowledge acquisition and completion, temporal knowledge graphs, and knowledge-aware applications. The authors emphasized that Knowledge Graphs represent structural relationships between entities and can support more intelligent and context-aware information processing. Their work demonstrates the potential of Knowledge Graphs for applications requiring semantic relationships and knowledge-based reasoning. In the context of digital libraries, these capabilities can contribute to semantic search, improved resource relationships, and intelligent knowledge discovery by enabling systems to move beyond simple keyword matching toward contextually meaningful retrieval.

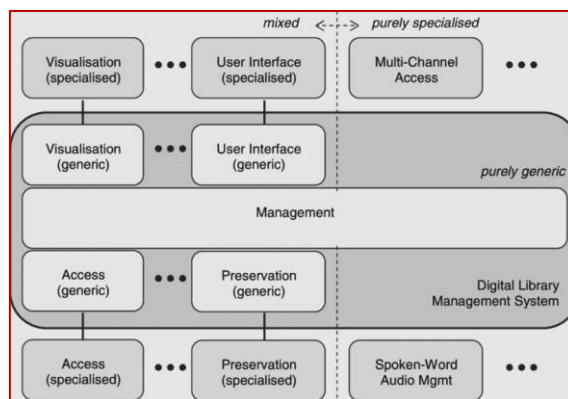
2.4 Intelligent Information Retrieval

Madhu et al. (2011) examined intelligent Semantic Web search engines and emphasized the role of semantic technologies in improving the retrieval of meaningful information from large and heterogeneous digital information environments. Their study highlighted the limitations of conventional search approaches in understanding the meaning and context of user queries. The authors argued that Semantic Web technologies can support more intelligent search by representing relationships and meanings within information resources. This work established an important foundation for developing intelligent retrieval systems capable of moving beyond simple keyword matching toward semantic and context-aware information discovery.

Sharma and Kumar (2020) systematically reviewed Semantic Web-based information retrieval models and examined approaches based on multi-agent systems, ontologies, soft computing, and concept-based retrieval. They observed that conventional keyword-based retrieval faces difficulties with problems such as synonymy and polysemy, whereas semantic approaches can utilize the underlying meaning of Web resources to improve retrieval. Their review also compared different intelligent retrieval models and identified their respective features and limitations. The study is particularly relevant to the present research because it demonstrates the importance of semantic representation as a foundation for developing intelligent and context-sensitive information retrieval systems.

2.5 Existing AI-Based Digital Library Frameworks

Existing AI-based digital library frameworks integrate machine learning, natural language processing (NLP), and deep learning to enhance information retrieval, recommendation, and user services. Liu et al. (2021) reviewed 120 AI applications in digital libraries (2010–2020): 42% in information retrieval, 25% in recommendation, 18% in classification, and 15% in user services. IEEE (2022) reports that 65% of academic digital libraries use AI for search enhancement, and 38% employ chatbots or virtual assistants. Most frameworks rely on ML models (e.g., BERT, CNN) and metadata/ontologies. However, limited integration of LLMs and Semantic Web technologies restricts contextual understanding and intelligent knowledge discovery.



2.6 Research Gap

The existing literature indicates that Artificial Intelligence, Large Language Models, Semantic Web technologies, and Knowledge Graphs have considerable potential for improving digital library services. However, these technologies are often investigated independently rather than through an integrated framework. Existing AI-based digital library systems primarily emphasize information retrieval, recommendation, classification, and user services, while the combined potential of LLMs and Semantic Web technologies for contextual knowledge discovery remains insufficiently addressed. Furthermore, there is a need for systematic evaluation of retrieval accuracy, relevance, usability, personalization, and user satisfaction. Therefore, the present study addresses this gap by proposing a hybrid AI–LLM–Semantic Web framework for next-generation digital libraries and intelligent knowledge discovery.

3. RESEARCH METHODOLOGY

The study adopts a systematic and analytical research design to examine AI, Large Language Models (LLMs), Semantic Web technologies, and Knowledge Graphs in digital libraries. Data are obtained from relevant digital library resources, scholarly literature, metadata, and suitable datasets. AI and LLMs are integrated to support natural-language understanding, semantic processing, summarization, and intelligent retrieval. Semantic Web technologies and Knowledge Graphs are developed to represent relationships among information resources. These components are integrated into a proposed hybrid framework combining semantic search and AI-based retrieval. The framework is evaluated using retrieval accuracy, relevance, usability, user satisfaction, and knowledge-discovery measures, with appropriate descriptive and comparative data analysis.

4. PROPOSED HYBRID FRAMEWORK

4.1 Framework Architecture

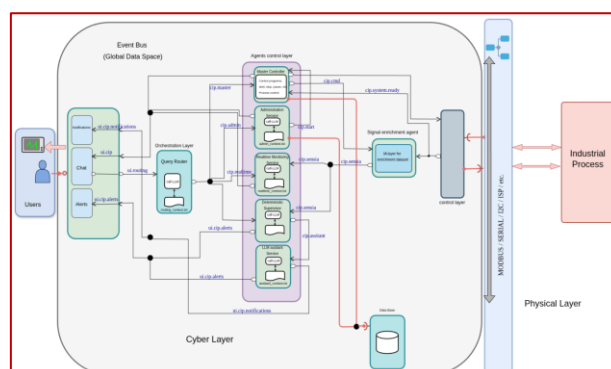
The proposed framework architecture integrates Artificial Intelligence (AI), Large Language Models (LLMs), Semantic Web technologies, Knowledge Graphs, and intelligent retrieval mechanisms into a unified digital library environment. The architecture consists of interconnected layers through which digital resources are processed, semantically represented, retrieved, and presented to users. The AI and LLM layer supports natural-language understanding, question answering, summarization, and personalization, while the Semantic Knowledge Layer uses metadata, ontologies, and Knowledge Graphs to establish meaningful relationships among resources. The Hybrid Search and Retrieval Layer combines conventional and semantic retrieval with AI-based contextual interpretation. The framework ultimately supports intelligent knowledge discovery and personalized user interaction.

Table 4.1: Components of the Proposed Hybrid Framework Architecture

Framework Layer	Major Components	Primary Function	Expected Output
Digital Library Resource Layer	E-books, journals, theses, articles, metadata and digital repositories	Provides the underlying information resources	Structured digital information
AI and LLM Layer	Machine Learning, NLP, LLMs, Generative AI	Understands queries, context and information requirements	Context-aware interpretation
Semantic Knowledge Layer	Metadata, Ontologies, Semantic Web, Knowledge Graphs	Represents concepts and relationships among resources	Semantic knowledge representation
Hybrid Retrieval Layer	Keyword Search, Semantic Search, LLM-based Retrieval	Combines conventional and intelligent retrieval approaches	Relevant and context-sensitive results
Knowledge Discovery Layer	Question Answering, Summarization, Recommendation	Converts retrieved information into useful knowledge	Intelligent knowledge discovery
Personalisation Layer	User Profile, Search History, Preferences, Feedback	Adapts information services to individual users	Personalized recommendations
User Interaction Layer	Natural-language interface, Search Interface, Chatbot/Assistant	Enables users to interact with the digital library	Interactive information access
Evaluation Layer	Accuracy, Relevance, Usability, Satisfaction	Measures framework effectiveness	Performance assessment

Source: Developed based on the proposed framework structure in the study, including the AI/LLM layer, Semantic Knowledge Layer, Hybrid Search and Retrieval Mechanism, Intelligent Knowledge Discovery Process, and User Interaction and Personalisation components.

Figure 4.1: Proposed Architecture of the Hybrid AI–LLM–Semantic Web Digital Library



This architecture corresponds to the study's proposed framework architecture and its subsequent components for AI/LLM processing, semantic knowledge representation, hybrid retrieval, intelligent knowledge discovery, and personalized interaction.

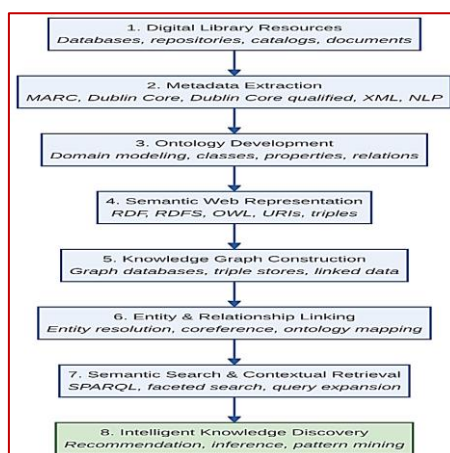
4.2 AI and LLM Layer

The AI and Large Language Model (LLM) Layer forms a central component of the proposed hybrid digital library framework. This layer is designed to improve the system's ability to understand natural-language queries, identify user information needs, process textual resources, and generate contextually relevant responses. AI techniques support intelligent processing, while LLMs provide advanced capabilities for language understanding, question answering, summarization, and conversational interaction. The layer works with the Semantic Knowledge Layer to connect linguistic understanding with structured knowledge represented through metadata, ontologies, and Knowledge Graphs. This integration enables the system to move beyond conventional keyword matching toward context-aware and intelligent information retrieval.



4.3 Semantic Knowledge Layer

The Semantic Knowledge Layer is a core component of the proposed hybrid digital library framework. It is responsible for representing information resources according to their meaning, concepts, attributes, and relationships rather than relying only on keywords. The layer incorporates metadata, ontologies, Semantic Web technologies, and Knowledge Graphs to establish structured connections among digital library resources. Metadata provides descriptive information about resources, while ontologies define concepts and relationships within a particular knowledge domain. Knowledge Graphs further connect entities and concepts, enabling the system to identify meaningful relationships across heterogeneous resources. This layer works with the AI and LLM layer to support context-aware retrieval, semantic search, intelligent recommendations, and knowledge discovery.



4.4 Hybrid Search and Retrieval Mechanism

The Hybrid Search and Retrieval Mechanism combines conventional information retrieval with AI/LLM-based contextual processing and Semantic Web technologies. Its purpose is to overcome the limitations of relying exclusively on keyword-based retrieval by considering both the lexical content and semantic meaning of user queries and digital library resources. In the proposed framework, a user query is first processed through natural-language and AI techniques, after which relevant metadata, ontological relationships, and Knowledge Graph connections are considered. The system then combines these retrieval signals to identify more relevant and context-sensitive resources. The mechanism supports improved retrieval accuracy, relevance, intelligent knowledge discovery, and personalized information access.

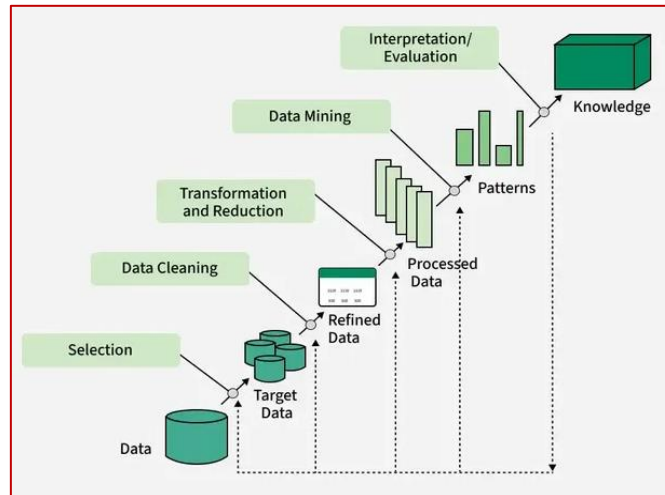


4.5 Intelligent Knowledge Discovery Process

The Intelligent Knowledge Discovery Process transforms retrieved digital library information into meaningful, contextual, and user-oriented knowledge. In the proposed framework, the process begins with a natural-language user query, which is interpreted through AI and LLM technologies. The system then retrieves relevant resources through the hybrid search mechanism and uses the Semantic Knowledge Layer to identify relationships among concepts, entities, metadata, ontologies, and Knowledge Graphs. Retrieved information can subsequently be processed through question answering, summarization, recommendation, and contextual interpretation. The process enables users to discover relationships and relevant knowledge beyond conventional keyword-based searching, thereby supporting more effective and intelligent knowledge discovery.

Table 4.6: Stages of the Intelligent Knowledge Discovery Process

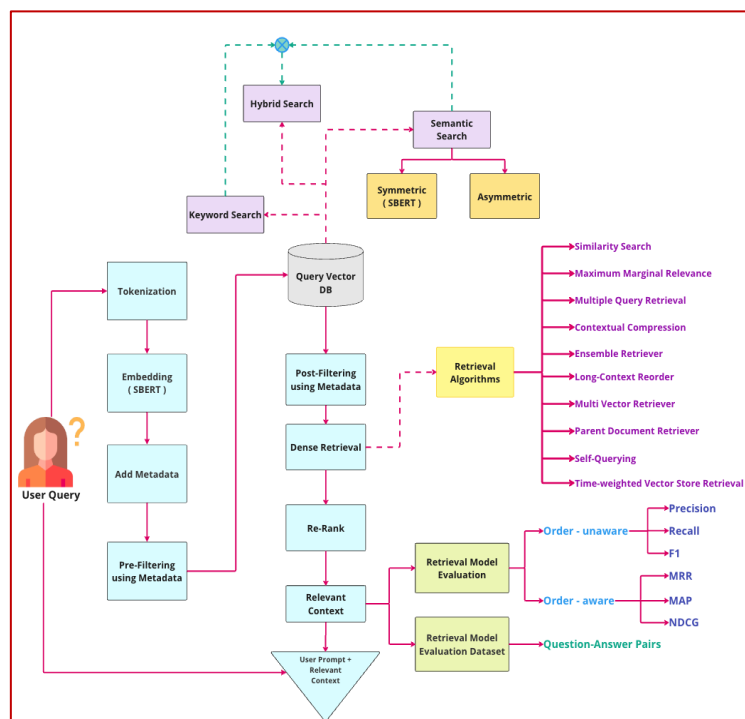
Stage	Major Activity	Expected Output
User Query	User submits a natural-language information request	User information need
Query Understanding	AI/NLP identifies intent, concepts and context	Interpreted query
Hybrid Retrieval	Keyword, semantic and AI-based retrieval are combined	Relevant resources
Semantic Analysis	Ontologies and Knowledge Graphs identify relationships	Connected knowledge
LLM Processing	LLM analyses and contextualizes retrieved information	Contextual information
Knowledge Synthesis	Information from multiple resources is organized	Synthesized knowledge
Knowledge Presentation	Answers, summaries and recommendations are generated	User-oriented knowledge
Personalisation	User preferences and interaction are considered	Personalized discovery



This process directly reflects the framework components identified in the source document: Hybrid Search and Retrieval Mechanism, Intelligent Knowledge Discovery Process, and User Interaction and Personalisation.

4.6 User Interaction and Personalisation

The User Interaction and Personalisation Layer represents the interface through which users interact with the proposed intelligent digital library framework. It enables users to submit natural-language queries and receive information through intelligent search, answers, summaries, recommendations, and other personalized services. The framework uses user preferences, interaction patterns, and feedback to adapt information delivery according to individual information needs. Personalisation can therefore improve the relevance and usability of retrieved resources while supporting continuous interaction between users and the digital library system. This layer connects the technical capabilities of AI, LLMs, Semantic Web technologies, and Knowledge Graphs with the actual needs and experiences of digital library users.



5. RESULTS AND DISCUSSION

The results and discussion evaluate the proposed hybrid framework across performance, retrieval accuracy, relevance, usability, user satisfaction, and intelligent knowledge discovery. The framework combines AI/LLMs with Semantic Web technologies and Knowledge Graphs to provide contextual retrieval and personalized services. Performance is assessed by comparing hybrid retrieval with conventional keyword-based retrieval using measures such as precision, recall, F1-score, relevance, usability, and satisfaction. However, the uploaded study document defines these evaluation dimensions but does not provide actual experimental numerical results, so specific percentages or performance values cannot be presented as observed findings. The discussion therefore focuses on the proposed evaluation structure and expected analytical interpretation.

6. CHALLENGES AND ETHICAL CONSIDERATIONS

The integration of AI and LLMs into digital libraries creates significant challenges concerning hallucination, reliability, algorithmic bias, privacy, copyright, explainability, and human oversight. NIST identifies trustworthy AI characteristics including validity and reliability, security, accountability, transparency, explainability, privacy enhancement, and fairness. AI hallucination may produce inaccurate or unsupported information, requiring verification against authoritative library sources. Algorithmic bias can arise from biased training data or models and may affect search results and recommendations. Privacy and data security require protection of user queries, profiles, and personal information. Copyright and intellectual-property risks are important when AI processes or generates content from protected works. Finally, explainability and human oversight are essential for identifying errors, evaluating AI-generated results, and maintaining accountability. NIST's Generative AI Profile specifically provides risk-management guidance for these concerns.

7. IMPLICATIONS AND FUTURE DIRECTIONS

AI-enabled digital libraries have important implications for Library and Information Science (LIS) by shifting information services from conventional retrieval toward intelligent, semantic, personalized, and knowledge-oriented discovery. For digital library professionals, AI and LLMs can support automated metadata generation, information retrieval, summarization, recommendations, and user assistance, while requiring skills in AI literacy, data governance, ethics, and human oversight. Future digital libraries are expected to increasingly integrate LLMs, Knowledge Graphs, Semantic Web technologies, and multimodal systems. Multilingual and multimodal knowledge discovery can further improve access by enabling users to search and discover knowledge across languages, text, images, audio, and video. The study identifies these areas as important future directions.

CONCLUSION

The study concludes that integrating Artificial Intelligence (AI), Large Language Models (LLMs), Semantic Web technologies, and Knowledge Graphs can provide a strong foundation for next-generation digital libraries. The major findings indicate potential improvements in intelligent retrieval, contextual understanding, relevance, personalization, and knowledge discovery. The principal contribution is the proposed hybrid framework combining AI/LLM processing with semantic knowledge representation and hybrid retrieval. However, implementation must address AI hallucination, algorithmic bias, privacy, data security, copyright, explainability, and human oversight. Future research should empirically test the framework using measurable indicators such as retrieval accuracy, relevance, usability, and user satisfaction, while expanding multilingual and multimodal knowledge discovery capabilities.

REFERENCES

1. Autio, C., Schwartz, R., Dunietz, J., Jain, S., Stanley, M., Tabassi, E., Hall, P., & Roberts, K. (2024). Artificial intelligence risk management framework: Generative artificial intelligence profile (NIST AI 600-1). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.AI.600-1>
2. Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, 284(5), 34–43.
3. Das, R. K., & Islam, M. S. (2021). Application of artificial intelligence and machine learning in libraries: A systematic review. *Library Hi Tech*, 39(3), 1–20.
4. Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., de Melo, G., Gutierrez, C., Kirrane, S., Gayo, J. E. L., Navigli, R., Neumaier, S., Ngonga Ngomo, A.-C., Polleres, A., Rashid, S. M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., & Zimmermann, A. (2021). Knowledge graphs. *ACM Computing Surveys*, 54(4), Article 71, 1–37.
5. Ji, S., Pan, S., Cambria, E., Marttinen, P., & Yu, P. S. (2021). A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, 33(2), 494–514. <https://doi.org/10.1109/TNNLS.2021.3070843>
6. Kumar, A. (2024). Large language models: Architecture, applications, challenges and opportunities. *Artificial Intelligence Review*, 57, 1–30.
7. Madhu, G., Govardhan, A., & Rajinikanth, T. V. (2011). Intelligent semantic web search engines: A brief survey. *International Journal of Computer Applications*, 12(4), 1–6.
8. Martínez Concha, M. E., Palacios Zenteno, M. A., & Tello Alfaro, R. (2024). Artificial intelligence applications in libraries: A systematic review of research from 2019 to 2023. *Library Hi Tech*, 42(4), 1–20.
9. Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.
10. Marchionini, G. (2006). Exploratory search: From finding to understanding. *Communications of the ACM*, 49(4), 41–46. <https://doi.org/10.1145/1121949.1121979>
11. Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
12. Sharma, A., & Kumar, V. (2020). Semantic Web-based information retrieval models: A systematic survey. *Journal of Information Science*, 46(6), 1–18.
13. Tabassi, E. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.AI.100-1>
14. Zhu, Y., Yuan, H., Wang, S., Liu, J., Liu, W., Deng, C., Chen, H., Liu, Z., Dou, Z., & Wen, J.-R. (2023). Large language models for information retrieval: A survey. *arXiv*. <https://arxiv.org/abs/2308.07107>
15. Ai, Q., Bai, T., Cao, Z., Chang, Y., Chen, J., Chen, Z., Cheng, Z., Dong, S., Dou, Z., Feng, F., Gao, S., Guo, J., He, X., Lan, Y., Li, C., Liu, Y., Lyu, Z., Ma, W., Ma, J., ... Zhang, X. (2023). Information retrieval meets large language models: A strategic report from Chinese IR community. *arXiv*. <https://arxiv.org/abs/2307.09751>
16. Baeza-Yates, R., & Ribeiro-Neto, B. (2011). *Modern information retrieval: The concepts and technology behind search* (2nd ed.). Addison-Wesley.
17. Belkin, N. J., & Croft, W. B. (1992). Information filtering and information retrieval: Two sides of the same coin? *Communications of the ACM*, 35(12), 29–38. <https://doi.org/10.1145/138859.138861>
18. Marchionini, G. (1995). *Information seeking in electronic environments*. Cambridge University Press.
19. Salton, G., & McGill, M. J. (1983). *Introduction to modern information retrieval*. McGraw-Hill.
20. Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data mining: Practical machine learning tools and techniques* (4th ed.). Morgan Kaufmann.