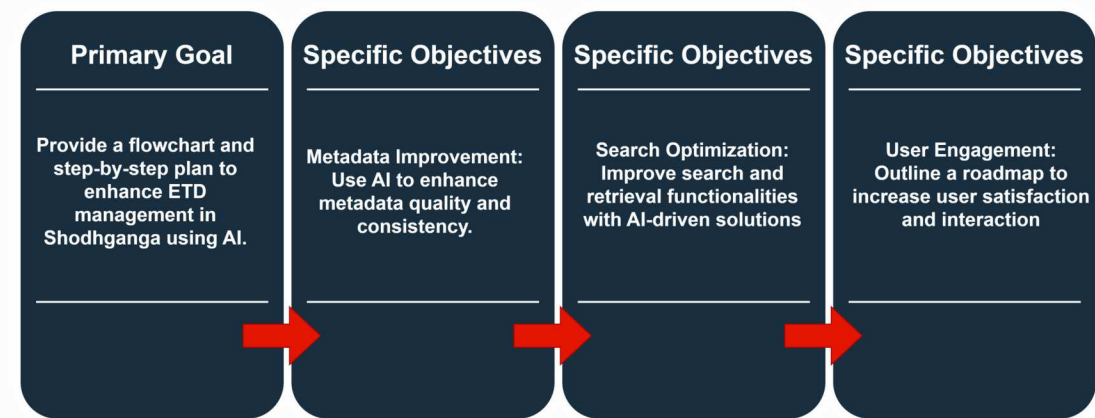


Introduction

The implementation plan for these AI-driven solutions is divided into three phases: Research and Development, Testing as prototype, and Full-Scale Deployment. During the Research and Development phase, a detailed literature review will identify best practices in AI-driven ETD management, and current management practices in Shodhganga will be analyzed to pinpoint areas for improvement. AI models for metadata creation, search enhancement, and recommendations will be developed and tested. The Pilot Testing phase will involve implementing these solutions on a subset of ETDs, collecting user feedback, and refining the algorithms based on initial results. Finally, in the Full-Scale Deployment phase, the AI solutions will be rolled out across the entire Shodhganga repository, with continuous monitoring and iterative improvements to ensure optimal performance and adaptation to new data.

Objectives



Methodology

This project's initial phase conducts a comprehensive literature review to identify best practices in AI-driven ETD management, specifically aimed at enhancing Shodhganga. By analyzing existing research, the review highlights successful AI applications in metadata improvement, search optimization, and user engagement.g phases—such as pilot testing and full-scale deployment—for future exploration.

Figure 1 Shodhganga Thesis Submission Process.

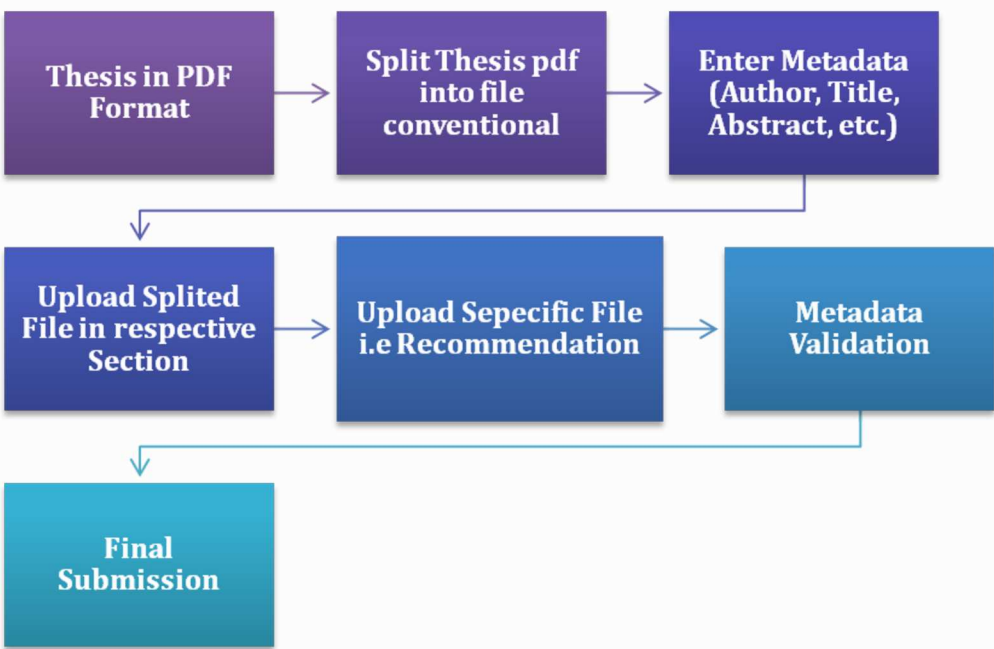
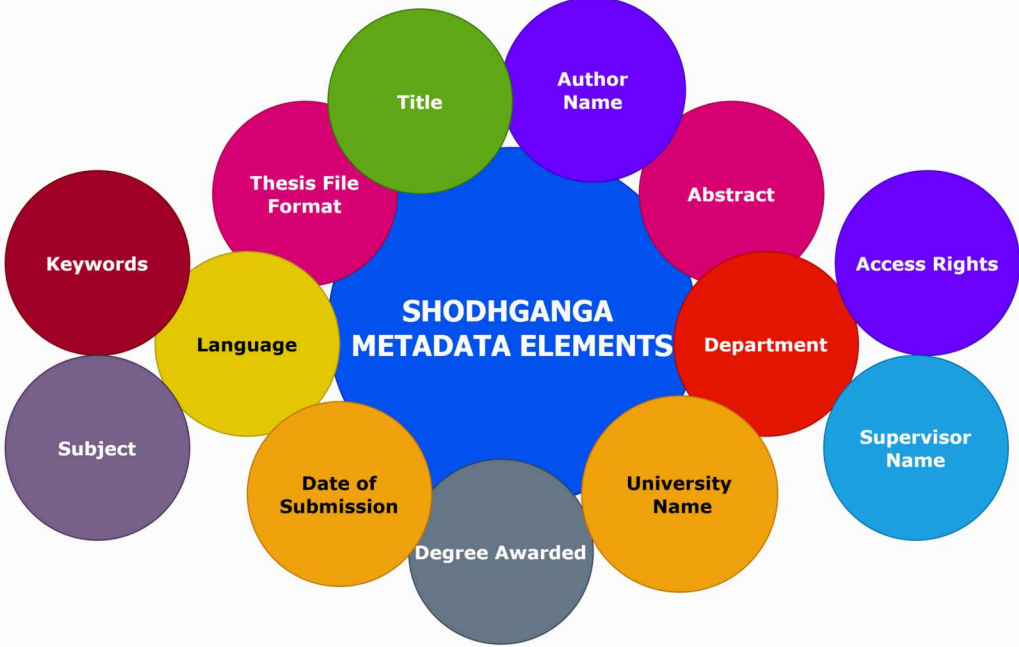


Figure 2 Shodhganga Thesis Metadata.



1. Illustrate AI-driven solutions for improving metadata quality and consistency

Figure:- 3AI-driven ETD Management Process.

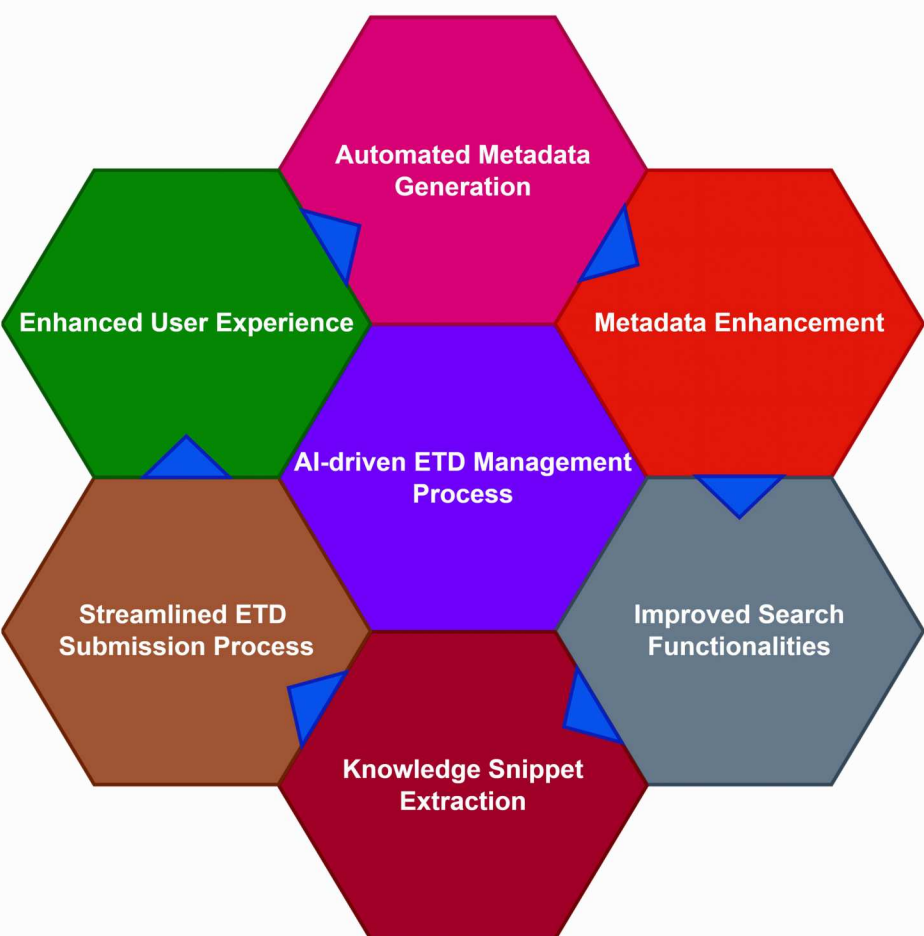
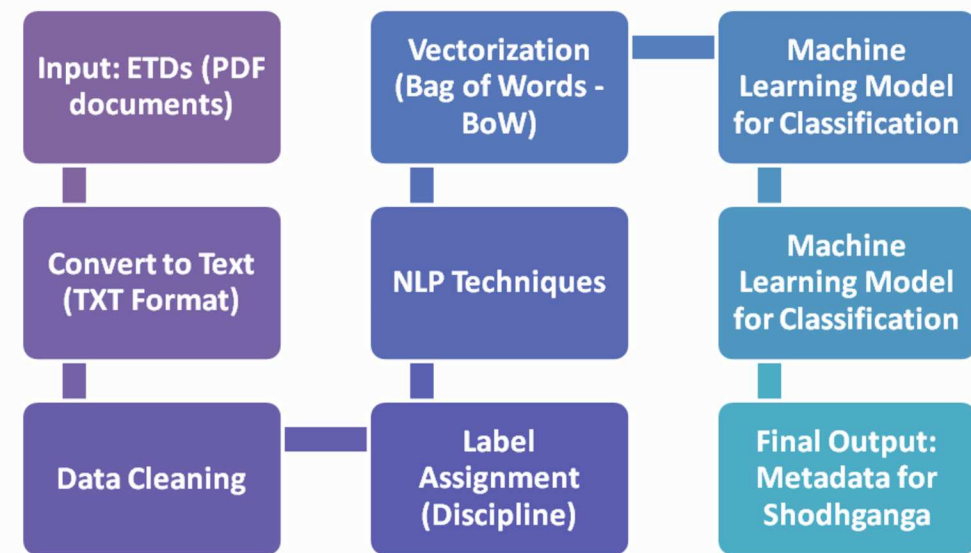


Figure : Flowchart of AI-driven ETD Management Process

AI-driven techniques can streamline processes in ETD repositories like Shodhganga by automating tasks such as metadata extraction, content classification, and personalized recommendations. These tools improve metadata accuracy, enhance search functionalities using natural language processing (NLP), and provide personalized recommendations based on user behavior and preferences (Singh & Varma, 2019).

Figure:- 4 Automated Metadata Extraction Flowchart



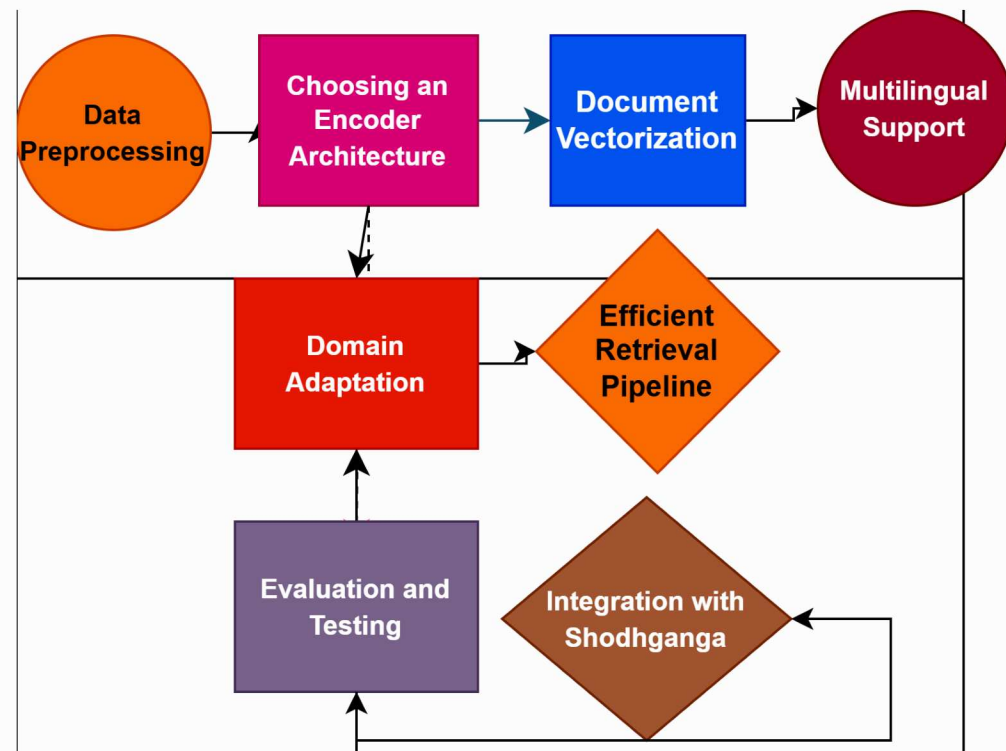
The figure from the paper "Metadata Extraction of RFIs Using Natural Language Processing" provides a clear workflow for automatic metadata extraction. This figure illustrates how the RFI documents undergo processing through multiple steps such as error detection, correction, and validation. In the Shodhganga context, a similar approach is used for ETDs (Electronic Theses and Dissertations), where the workflow involves.

- Automatic metadata extraction using NLP techniques.
- Classification of metadata fields based on predefined categories.
- Error correction and validation using machine learning models.

By employing AI technologies for metadata validation and enhancement, as shown in Figure of the cited paper, the Shodhganga automated the metadata.

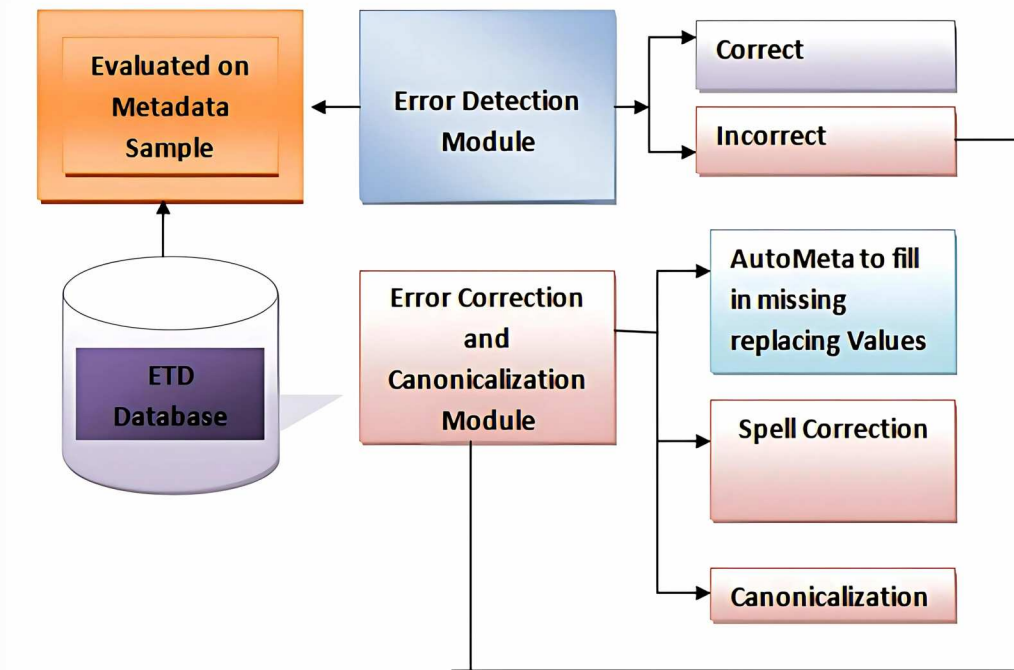
2. Outline steps to enhance search and retrieval functionalities

Figure:- 5 AI-driven Enhance search and retrieval functionalities.



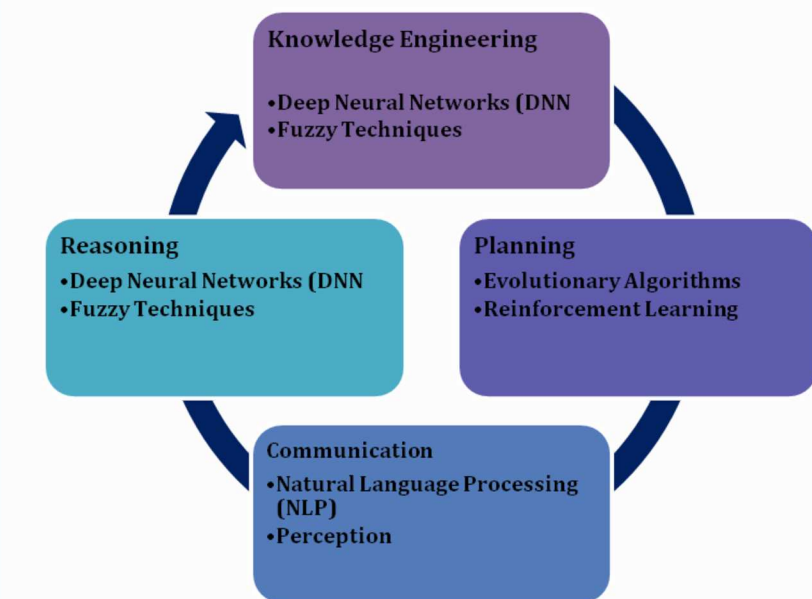
The semantic search model for Shodhganga can be implemented using bi-encoders and cross-encoders to enhance the retrieval of theses and dissertations. The system involves converting documents into structured vectors, storing them in a vector search engine like FAISS, and applying multilingual models like mBERT to support multiple Indian languages. A two-stage retrieval process is employed: bi-encoders for fast retrieval and cross-encoders for accurate re-ranking.(Bringmann & Zhukova, 2024).

Figure:- 6 Flowchart improving ETD metadata quality



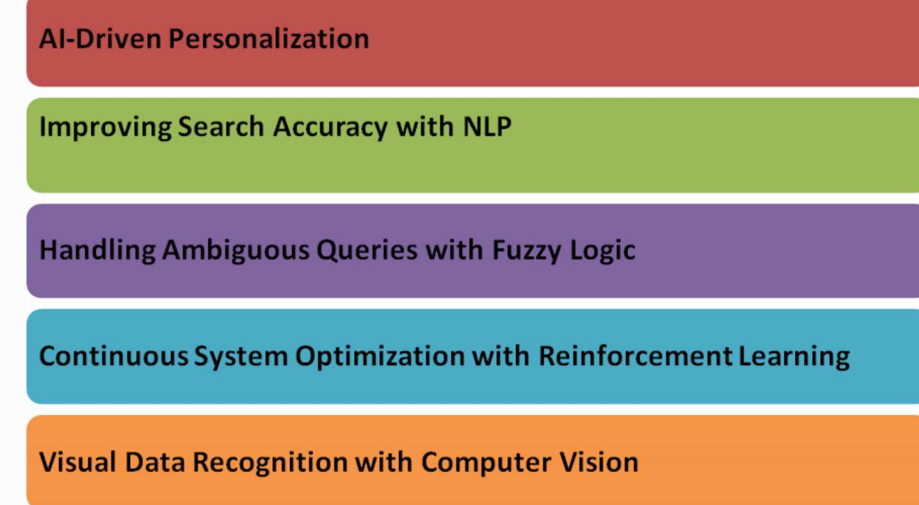
The system described can be applied to repositories like Shodhganga, where metadata quality is crucial for enhancing the accessibility and discoverability of Electronic Theses and Dissertations (ETDs). In this approach, metadata from the Shodhganga database would undergo error detection. If errors are found, the system uses tools like AutoMeta to fill missing values and applies spell correction and canonicalization to standardize and correct the metadata. This process ensures that the metadata stored in Shodhganga is accurate and consistently formatted, improving its overall quality and usability (Choudhury et al., n.d.).

Figure:- 7 Figure Interpretation for Recommendation System



AI techniques such as deep neural networks (DNNs), fuzzy logic, and reinforcement learning can significantly improve Shodhganga's recommendation system. As Salehi et al. (2020) explain, DNNs analyze user behavior to offer personalized recommendations, while fuzzy logic handles ambiguous search queries to ensure relevant results. Reinforcement learning optimizes the system based on user feedback. These AI-driven methods will enhance search accuracy and user engagement in the Shodhganga ETD repository.

Figure:- 8 Roadmap for Increasing User Engagement and Satisfaction in Shodhganga ETD Repository



This figure illustrates how AI techniques like deep neural networks (DNNs) can enhance Shodhganga by providing personalized recommendations, while natural language processing (NLP) improves its ability to interpret complex user queries. These advancements significantly boost user engagement and satisfaction by optimizing search accuracy and relevance (Salehi et al., 2020).

Conclusion

AI integration into Shodhganga will address critical challenges, enhancing metadata quality search functionality, and user experience. This will position Shodhganga as an advanced academic repository, expanding its reach and value for the scholarly community.

Outcomes

Integrating AI into Shodhganga's ETD management will streamline processes, reduce manual workload, and enhance operational efficiency. Key outcomes include:

- Enhanced metadata quality and consistency.
- Improved search functionalities and personalized recommendations.
- Expanded integration with external databases, leading to broader research access.
- Increased user engagement and satisfaction, making Shodhganga a more valuable academic resource.

References

- Bringmann, A., & Zhukova, A. (2024). Domain adaptation of multilingual semantic search: Literature review. Georg-August University Göttingen. <https://arxiv.org/abs/2402.02932>
- Choudhury, M. H., Salsabil, L., Jayanetti, H. R., Wu, J., Ingram, W. A., & Fox, E. A. (n.d.). MetaEnhance: Metadata quality improvement for electronic theses and dissertations of university libraries. Old Dominion University and Virginia Polytechnic Institute and State University.
- Kuniyoshi, S. K., Schlichting, H. E., & Fujii, Y. (2023). Domain adaptation of multilingual semantic search. arXiv. <https://arxiv.org/abs/2303.17661>
- Ozoğlu, C., & Ergen, M. (2020). Metadata extraction of RFIs using natural language processing. Semantic Scholar. <https://www.semanticscholar.org/paper/Metadata-Extraction-of-RFIs-Using-Natural-Language-%C3%96zo%C4%9Ful-Ergen/9faa69292bdea00bc03c76a61b454494598780e9>
- Salehi, M., Pouriyeh, S. A., Seno, S. A. H., Safavi, S. M., Shariatmadari, S., & Niknam, M. (2020). Artificial intelligence in recommender systems. Journal of Computer Networks and Communications, 2020, 1–23. https://www.researchgate.net/publication/346501646_Artificial_intelligence_in_recommender_systems
- Singh, M. K., & Varma, N. (2019). Artificial intelligence in libraries: A review. Library Philosophy and Practice (e-Journal). <https://digitalcommons.unl.edu/libphilprac/3511>

AUTHORS :-

- Vivek Ranjan¹(Assistant Librarian, Silver Oak University, INFLIBNET Centre)
- Nabajit Saikia²(Software Developer, Shodhganga, INFLIBNET Centre),
- Manoj Kumar K³(Scientist- F(CS),Shodhganga, INFLIBNET Centre)
- Surbhi⁴(Scientist-C(LS),Shodhganga, INFLIBNET Centre)