Developing a Smart Archival Assistant with conversational features and linguistic abilities: the Ask_ArchiLab initiative

Basma Makhlouf Shabou
Information Science Dept.
Geneva School of Business Administration University of Applied Sciences and Arts of Western Switzerland (HES-SO)
Geneva, Switzerland
https://orcid.org/0000-0003-0980-0517
basma.makhlouf-shabou@hesge.ch

Lamia Friha
R&D Cell, Research and Information Science, DiSTIC, University of Geneva
Geneva, Switzerland
https://orcid.org/0000-0003-3134-1070
Lamia.friha@unige.ch

Wassila Ramli
Corporate analytics and Digital project officer Contexa SA, University of Geneva
Geneva, Switzerland
https://orcid.org/0009-0009-0454-9387
wassila.ramli@unige.ch

***Abstract*— This article describes Ask_ArchiLab project, a recent project conducted at ArchiLab_Geneva School of Business Administration. The project aims to modernize archival practices using conversational AI.**

It addresses challenges in digital archiving through a multi-agent system that enables fast, contextual queries. The current focus is on delivering professional archival knowledge via semantic technologies like RDF using advanced RAG techniques. The project fosters international collaboration to enhance access and usability in archival science.

***Keywords*— Generative AI, RDF, Open Linked Data, GraphRAG, Agentics, Knowledge engineering, advanced archival knowledge, PIAf, ICA**

I. Introduction

Digital transformation has fundamentally changed how archival knowledge is managed, accessed, and shared in our modern era [1, 2, 3]. This project introduces the "Smart Archival Assistant," an advanced conversational AI system designed to revolutionize archival practices while addressing current challenges in organizing and retrieving vast amounts of digital information. The system aims to serve archival professionals and trainers through an intuitive interface that combines traditional archival principles with cutting-edge technology.

The Smart Archival Assistant leverages natural language processing, generative AI methods and advanced knowledge engineering techniques to facilitate natural conversational interactions with users, enabling efficient access to archival resources and knowledge discovery. This innovative solution primarily serves professional archivists, records managers, trainers, educators, and research scholars, while also accommodating the needs of public researchers and cultural heritage institutions. Through its advanced capabilities, the system supports various aspects of archival work, from resource location to knowledge synthesis and pattern recognition.

By implementing state-of-the-art technologies, the assistant offers personalized experiences that adapt to different user expertise levels while maintaining high standards of data privacy and archival integrity. The system prioritizes professional efficiency by streamlining workflows and enhancing decision-making processes in archival management. It also plays a crucial role in knowledge enhancement by facilitating expertise transfer between professionals and enabling the discovery of new connections within archival collections.

The project acknowledges certain boundaries in its scope, focusing primarily on professional archival content and practices while ensuring integration of capabilities with existing systems. Ethical considerations remain at the forefront, emphasizing data security, transparent AI decision-making, and inclusive information access. The Smart Archival Assistant represents a significant advancement in modernizing archival practices while preserving fundamental archival science principles, positioning itself as an indispensable tool in the evolving landscape of digital archives management.

This innovative approach addresses critical needs in the archival community, from streamlining access to vast repositories to supporting professional development initiatives. The system's ability to combine traditional archival knowledge with advanced technology creates a more accessible, efficient, and user-friendly experience for all stakeholders involved in archival work and research.

II. Problem and Needs

A. Problems

Contemporary archival management faces unprecedented challenges due to the exponential growth of digital records and complex data formats. This digital transformation creates significant barriers in achieving comprehensive, accessible, and open knowledge management. The challenges manifest throughout the archival process, affecting both operational procedures and educational initiatives [1, 2, 3, 4, 5].

These obstacles can be divided into several categories, including financial, technical, cognitive, linguistic, academic, professional, and sociocultural challenges.

Financially, many institutions face limited budgets for purchasing international standards, insufficient funding for training and professional development, and cost constraints that hinder the adoption of modern archival solutions.

On the technical side, organizations often struggle with tracking and implementing emerging archival methodologies, have limited access to advanced tools and technologies, and encounter difficulties managing diverse digital formats and integrating with legacy systems.

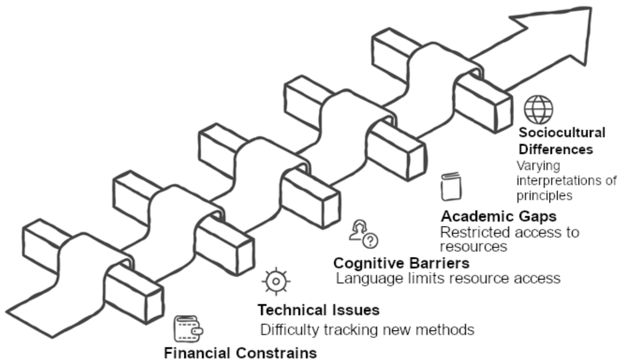

Fig. 1 Archival Process Hindered by Key Barriers

Cognitive and linguistic barriers also pose significant challenges, as language differences can restrict access to international resources, and the complexity of technical terminology across multiple languages complicates understanding. Additionally, there is limited accessibility to multilingual archival materials, and translation difficulties can impede international collaboration.

From an academic and professional perspective, restricted access to scholarly resources, limited opportunities for ongoing development, gaps between theoretical knowledge and practical application, and hurdles in applying international standards further complicate archival work.

Sociocultural factors, such as cultural differences in archival practices, varying interpretations of archival principles, regional variations in documentation methods, and diverse user needs and expectations, also contribute to these challenges. These latest significantly impact on archival professionals, students, and trainers, creating a pressing need for innovative solutions in knowledge access and management. The complexity of these problems requires a comprehensive approach that addresses both immediate practical needs and long-term strategic objectives.

B. Needs (Personas)

The following cases and expressed needs with different roles and personas. Six illustrative personas are presented in following paragraphs:

1) History Enthusiast

Meet Sarah, a history enthusiast with a passion for uncovering hidden tales from the past. As Sarah delves into historical documents and archives, she often finds herself overwhelmed by the sheer volume of information. A conversational chatbot equipped with archival knowledge could help Sarah navigate these archives efficiently, providing her with quick answers and context to enrich her historical research.

2) Small Business Owner

John runs a small antique shop filled with fascinating relics, but organizing his collection and keeping track of provenance information is a challenge. A chatbot tailored to archival best practices could assist John in cataloging his inventory, preserving crucial details for each item, and even recommending strategies for proper conservation.

3) Student Researcher

Emily, a university student specializing in anthropology, is amid a research project requiring extensive exploration of online archives. With a conversational chatbot, Emily could streamline her search process, receive guidance on refining her queries, and discover new avenues for sourcing valuable primary documents.

4) Genealogy Enthusiast

David is deeply invested in tracing his family history but often finds himself stuck at roadblocks when piecing together his lineage. By utilizing a chatbot that understands archival principles, David could receive tailored tips on genealogical research, access curated databases, and connect with fellow enthusiasts to exchange insights and resources.

5) Archival Science Professor

Dr. Patel, an esteemed professor in archival sciences, is dedicated to educating future archivists on the best practices in preserving historical records. A conversational chatbot designed to assist Dr. Patel could provide real-time examples, case studies, and interactive exercises to enhance classroom learning and mentor students in their archival journey. The same facility may suggest possibly exam questions to help Dr. Patel in his job.

6) Entry-level Archivist

Sophie, a recent graduate embarking on her career as an archivist, is eager to apply her knowledge but faces challenges in navigating complex archival systems. A chatbot customized for new professionals like Sophie could offer guidance on archival standards, facilitate networking opportunities within the field, and provide helpful resources for professional development.

In a landscape where archival expertise is essential for various individuals, from seasoned professionals to budding enthusiasts, the integration of a conversational chatbot revolutionizes the accessibility and usability of archival knowledge. These personas reflect a diverse spectrum of users who could benefit significantly from such technological support in their archival endeavors.

In this project, we propose to build a conversational agent that leverages natural language processing, deep learning techniques to interact with the user and returns the most appropriate answer. Moreover, the agent will learn continuously to provide personalized answers and to make recommendations and user assistance.

III. Objectives

The solution requires integrating advanced generative AI technologies with semantic knowledge representation systems. The research goal of this project is to:

- Develop advanced conversational features based on AI latest technologies.
- Ensure fast and accurate search within archives and archival resources.
- Provide contextual and relevant responses to archival queries.
- Open archival knowledge to a wide range of publics.
- Enhance access to archived information for various users, including researchers, students, and the public.

A. Research Domains

Archival science is an interdisciplinary field of research that seeks to apply techniques and principles to manage and preserve archives. It also benefits from several research fields of **Artificial Intelligence (AI)**:

- Knowledge Representation and engineering
- Machine Learning
- Generative AI
- Natural Language Processing (NLP)
- Advanced Data Visualization

B. Core Project Objectives

- Develop a relevant **knowledge representation** of archival data and knowledge
- Develop a **conversational agent** on top of this knowledge representation
- Develop a distributed problem-solving approach trough **a multi-agent system**
- Provide advanced **user interaction** solutions
- **Test and evaluate** the developed conversational agent against the needs of the target audience

C. Project Complexity Factors

- **Multilingual** content management
- **Heterogeneous** data formats
- **Complex** metadata relationships
- **Diverse** user expertise levels
- **Cross-cultural** archival practices
- **Integration** with existing systems

D. Implementation Strategy

An interdisciplinary team combining Information Science experts from HEG-GE and AI specialists from the University of Geneva will address these challenges. Collaboration ensures balanced attention to both archival principles and technological innovation.

E. Value Proposition

- **Automated** processing of multilingual archives
- **Unified** access to heterogeneous formats
- **Intelligent** context-aware search
- **Personalized** user experiences
- **Cross-collection** knowledge discovery
- **Enhanced** accessibility for diverse users

IV. System Architecture

A. Initial Agent Framework

The schema (Fig. 1) illustrates the AskArchiLab system architecture with four main components:

1) Data Collection:

- Restricted access datasets (standards, reports, courses)
- Unrestricted access datasets (dictionaries, terminology, open portal, academic courses)
- All data flows into a DataWarehouse

2) Data Preparation and Processing:

- Sequential steps: ingestion, cleaning, transformation, standardization, enrichment, filtering, quality assurance, and labeling
- Each step ensures data quality and compatibility

3) Core System (AskArchiLab):

- User query processing through embedding model
- Triplestore for RDF representation
- Vector database for knowledge storage
- Retrieval system finding top k similar documents
- Context integration with query
- LLM for generating final responses
- Direct LLM answering for simpler queries

4) End users include professors, archivists, researchers, and students.

Key innovation: Dual-path architecture combining direct LLM responses with retrieved context-enhanced answers.

B. Distributed Artificial Intelligence

The Archival domain represents a complex problem space that naturally calls for a distributed problem-solving approach. As highlighted by Russell and Norvig (2020), "a distributed artificial intelligence (DAI) system consists of a group of intelligent agents that interact to solve problems that are beyond their individual capabilities. These distributed systems benefit from parallel computation and distributed data, and they can survive individual agent failures" [6]. This foundational concept perfectly matches the distributed characteristics inherent in archival work and knowledge management. This concept is empirically validated by Tsague et al. (2013), who demonstrated through their multi-agent modeling approach how electronic archiving systems inherently require distributed intelligence to handle "technical, security, legal, and organizational constraints" that span across multiple dimensions of the archival domain [7].

C. Distributed Nature of Archival Domain

The archival domain shows intrinsic distribution across multiple dimensions. From a geographical perspective, archives span across various institutions, countries, and cultural contexts, with each maintaining its own distinct collection policies and preservation methods. This physical distribution is enhanced by the distribution of knowledge itself, as expertise crosses multiple disciplines including history, preservation science, and digital technologies. Pellegrino et al. (2014) provide empirical evidence for this distributed nature through their research on autonomous digital preservation, demonstrating how "agents monitor and control the local environment and deal with preservation issues individuating obsolescent formats based on global parameters" [8]. Their findings confirm that preservation challenges inherently require distributed intelligence, as no single agent can effectively manage all aspects of digital preservation across diverse formats and contexts.

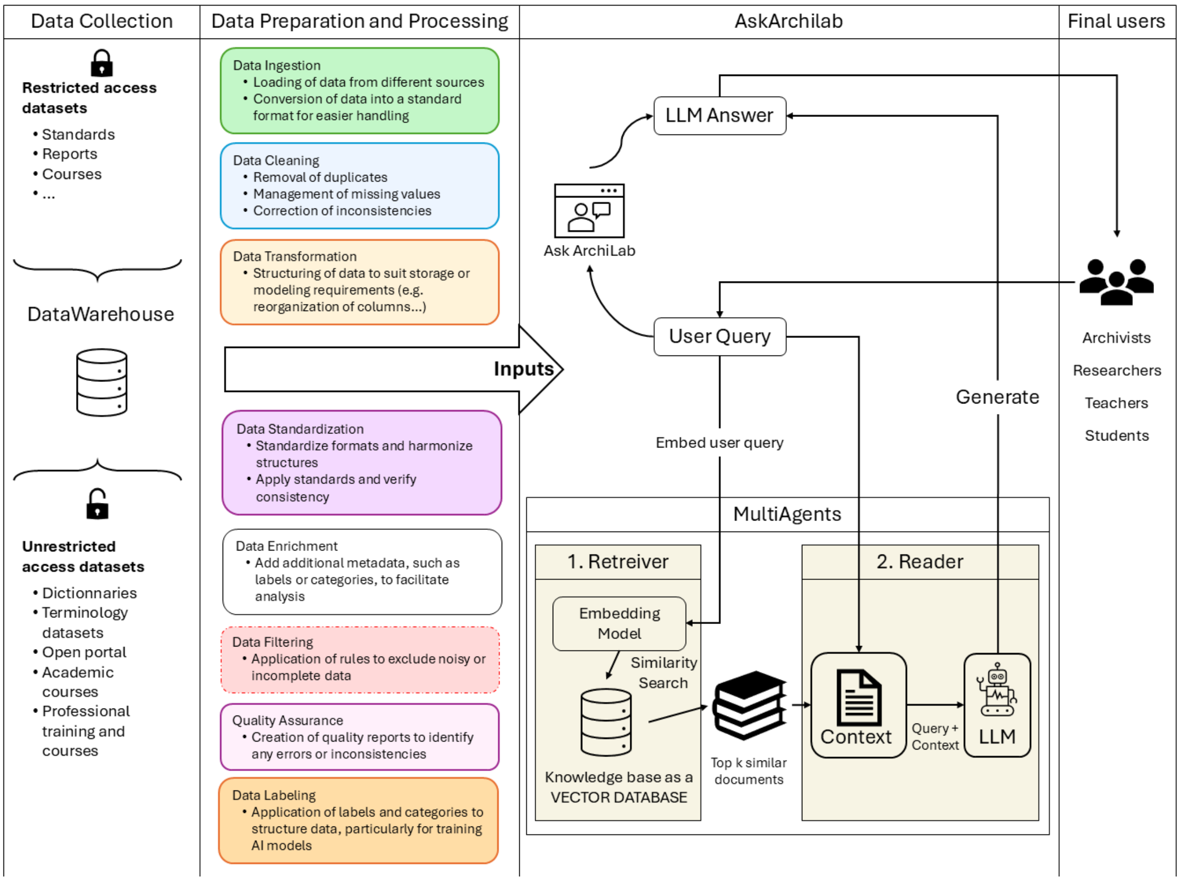

Fig. 2. Initial Agent Framework

Within the archival domain, information naturally distributes across various formats, from physical artifacts to digital collections, each demanding specific handling and expertise. This distribution pattern extends into classification systems, metadata schemas, and linguistic contexts, forming a rich yet complex ecosystem of interconnected knowledge [7]. The inherently parallel nature of archival work, where multiple processes must function simultaneously while maintaining coherence and consistency, further underlines this distributed character.

D. Towards a Multi-agent System

Building on this understanding of the domain's distributed nature, this project will expand to evolve our initial single conversational agent into a comprehensive multi-agent system. This transition is strongly supported by Aryal et al. (2024), whose recent research demonstrates the effectiveness of "domain-specific multi-AI agent system in identifying and bridging knowledge gaps." [9] Their comparative analysis of different multi-agent workflows provides concrete evidence that collaborative AI systems can effectively handle complex, interdisciplinary queries while maintaining system coherence.

The multi-agent framework will harness parallel processing capabilities, enabling simultaneous handling of various archival tasks while maintaining system coherence [8]. Aryal et al.'s findings specifically validate this approach, showing how their "platform aims to leverage the unique strengths and perspectives of each [agent]" through various workflow configurations that demonstrated superior performance in terms of efficiency and accuracy [9].

The system's design will emphasize collaborative problem-solving capabilities, enabling agents to share knowledge and coordinate actions across distributed components [7]. This mirrors the collaborative nature of archival work, where different expertise and resources must combine to achieve comprehensive solutions. The framework will incorporate mechanisms for dynamic adaptation to varying workloads and user needs, ensuring efficient resource utilization while maintaining high performance standards [8].

The effectiveness of this distributed approach depends on careful consideration of interoperability between different archival systems and standards, cultural and linguistic adaptability, and the ability to scale effectively with growing collections [7, 8]. Our research will investigate these aspects, exploring how the natural distribution in archival work can be effectively translated into a robust multi-agent system architecture that enhances rather than replaces existing archival practices.

This evolution from a single agent to a multi-agent system represents not just a technological advancement, but a fundamental alignment with the distributed nature of archival work itself [9]. The research will explore how this distributed intelligence can better serve the complex needs of archival management while maintaining system reliability and adaptability to changing requirements.

E. *Semantic AI*

The evolution of semantic technologies has significantly impacted how we manage and access knowledge in digital environments. In the context of archival systems, semantic AI approaches offer powerful capabilities for representing complex relationships, ensuring accurate information retrieval, and maintaining contextual understanding across diverse collections. This advancement in semantic processing is particularly relevant for archival science, where meaning and context are crucial for proper interpretation and use of archival materials.

1) *RDF: Fundamental Semantic Framework*

Resource Description Framework (RDF) has emerged as one of the most established semantic frameworks for knowledge representation. At its core, RDF provides a standardized model for data interchange on the Web, representing information through subject-predicate-object triples. This structure allows for sophisticated representation of archival relationships, hierarchies, and metadata. RDF's flexibility enables archives to create rich semantic networks where resources are uniquely identified through URI [1] s, relationships are explicitly defined, and complex queries become possible through SPARQL [2] endpoints. The framework's standardization by W3C has made it a cornerstone for linked data initiatives in cultural heritage institutions, facilitating interoperability and knowledge sharing across different archival systems.

2) *RAG: Advanced Techniques for Knowledge Retrieval and Generation*

Recent developments in AI have introduced Retrieval Augmented Generation (RAG) as an innovative technique for enhancing information access and generation. RAG represents a significant advancement by combining traditional retrieval systems with large language models to produce more accurate and contextually relevant responses. In archival applications, RAG and its variants (such as GraphRAG and Multi-Vector RAG) provide crucial capabilities for grounding AI responses in verified archival sources, reducing hallucination risks, and maintaining provenance of information. This approach is particularly valuable in archival contexts where accuracy and source verification are paramount. By integrating retrieval mechanisms with generative capabilities, RAG enables dynamic and intelligent access to archival knowledge while preserving the rigorous standards required for archival practice.

V. Practical Workflow Implementation

To operationalize the proposed retrieval-augmented architecture in Figure 1, a multi-layered workflow has been designed to ensure seamless interaction between user input, information retrieval, and language model processing. Each layer contributes to a specific function within the overall pipeline, forming a modular and scalable system suitable for archival data management and intelligent query answering.

As illustrated in Figure 2, the workflow is structured into six interdependent layers that process the query sequentially—from initial user interaction to the generation and delivery of a finalized, context-enriched response. This architecture ensures both technical robustness and semantic alignment, allowing the system to maintain archival accuracy while leveraging advanced natural language understanding and generative capabilities.

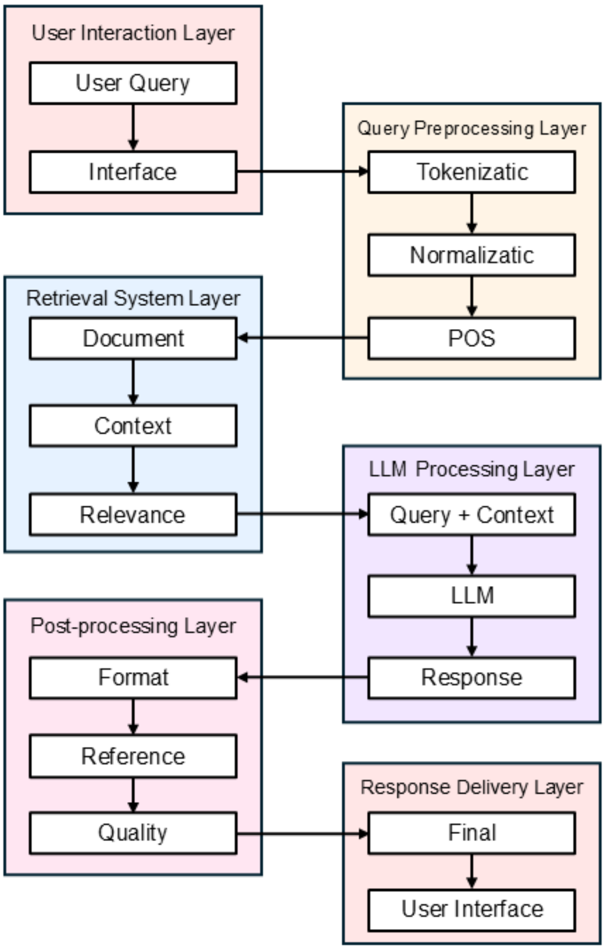

Fig. 3. Practical workflow implementation

A. *User Interaction Layer*

This first layer manages the user interface, capturing queries, and maintaining the conversational session context. Inputs are validated, logged, and forwarded to the backend. In line with modular RAG practices, the system may reformulate or disambiguate the query to improve retrieval accuracy.

B. *Layer 2: Query Preprocessing Layer*

The preprocessing layer tokenizes, normalizes, and enriches the raw query using part-of-speech (POS) tagging and named-entity extraction. It then converts the processed text into semantic embeddings that preserve contextual meaning. This ensures that the query is compatible with both keyword-based and semantic retrieval components.

[1] Universal Resource Identifier: Universal Resource identifiers in WWW

[2] SPARQL 1.1 Protocol

C. Layer 3: Hybrid Retrieval System Layer

This layer performs multi-strategy searches within the archival knowledge base to retrieve the most relevant contextual information. Dense vector search identifies semantically similar documents through embedding matching, while the RDF triplestore lookup retrieves interconnected triples and relationships. In parallel, sparse retrieval methods such as BM25 keyword scoring capture exact term matches to complement semantic retrieval. By combining dense and sparse strategies, the system maximizes contextual coverage and relevance. The top-ranked passages and RDF triples are then extracted, ranked, and prepared as input for contextual integration in the subsequent LLM processing stage.

D. Layer 4: LLM Processing Layer

In this core layer, the preprocessed query is integrated with the retrieved contextual information to generate a grounded response. Through contextual fusion and prompt engineering techniques, the LLM is provided with both the user query and the most relevant excerpts or RDF triples from the retrieval layer. This ensures that the generative process remains anchored in verified archival data. The LLM synthesizes this information to produce coherent, contextually accurate, and factually grounded answers. This dual-path architecture—combining direct LLM generation with retrieval-augmented context—enhances semantic consistency, reduces hallucinations, and improves overall response reliability.

E. Layer 5: Post-processing Layer

In this layer, the raw response generated by the LLM is refined to ensure clarity, factual accuracy, and compliance with archival quality standards. The process involves formatting the output, appending appropriate source citations, and verifying key claims against authoritative references. Validation mechanisms are applied to maintain consistency with archival standards, while readability and structural coherence are enhanced. Through these refinement steps, the final response is well-structured, properly referenced, and aligned with the expected archival and scholarly quality criteria.

F. Layer 6: Final Response Layer

The final layer packages the refined output for user delivery. It formats the response for display, logs the interaction, and may collect user feedback. A feedback loop supports continuous improvement of both retrieval and generation modules. The complete workflow is thus end-to-end: a query enters through the user interface and exits as a polished, context-enriched answer.

The multi-layered workflow described above demonstrates technically coherent and modular architecture capable of integrating retrieval and generative processes in a unified pipeline. Each layer communicates sequentially, forming a pipeline where output from one stage feeds into the next, allowing individual components to be updated or scaled independently. For example, the hybrid retrieval layer can be tuned or extended (e.g., adding new triple store queries) without altering the LLM layer. Each layer contributes to ensuring data consistency, factual grounding, and contextual enrichment, while maintaining interoperability across archival and linguistic resources. The systematic flow—from query preprocessing to response delivery—provides a robust foundation for scalable deployment and iterative optimization, mirroring state-of-the-art RAG pipelines to enable fast, accurate, and contextually rich responses to archival queries. This implementation not only operationalizes the conceptual model but also establishes a framework for empirical evaluation. The following section, Preliminary Results, presents an initial assessment of the system's performance, focusing on retrieval accuracy, response quality, and the overall effectiveness of the hybrid LLM–retrieval integration in the archival context.

VI. Preliminary Results

A. Case 1: ICA resource center

In our initial prototype, we tested the AskArchiLab agent on content from the ICA (International Council on Archives) Online Resource Centre. This resource centre hosts a wide array of archival materials – including toolkits, guides, manuals, standards, glossaries, directories, and bibliographies. We collected the available documents and training materials from the ICA site and applied a rigorous preprocessing pipeline: the text was cleaned, segmented, and mapped into RDF triples and embeddings. The result is a structured knowledge corpus representing much of the ICA's professional archival guidance (with provenance retained for traceability).

Using this ICA dataset, we conducted example query tests to evaluate the workflow. For instance, when asked about descriptions or standards (e.g. "What is the ICA standard for archival description?"), the system retrieved relevant ICA documents (such as ISAD(G) guidelines) and grounded its answer in those sources. The LLM then generated responses that accurately reflected the ICA material, benefiting from the context injection step. In practice, the agent was able to surface specific archival definitions and recommendations directly from the ICA content. These early experiments indicate that the contextual RAG approach works as intended: the LLM's answers are coherent and factually based on the archival corpus. In summary, the Case 1 results demonstrate the feasibility of our layered system – the agent can effectively leverage the ICA resource data to answer domain-specific queries with high relevance and accuracy [11].

The success of this initial test paves the way for expanding additional sources (e.g. the PIAF dataset) and for quantitative evaluation of retrieval accuracy in future work.

B. Case 2: PIAF data set

PIAF stands for the International Francophone Archival Portal. It is an important source for our ask_archilab project. It aims to feed the first level of the system. PIAF describes a set of training modules written by international experts and covering the entire archival management cycle [10]. Each module contains thematic online sections, additional pdfs documents, and a hierarchical organization that structures the whole. A rigorous preprocessing has been applied to this data in various formats. As part of a Bachelor thesis, a complete transformation of the PIAFs educational content was completed and produced a structured, cleaned and enriched dataset to serve within the ICA and the foundation for the Ask_archilab project. The developed pipeline successfully addresses the transformation. The resulting corpus is over 90% complete and maintains source traceability.

VII. Evaluation and Metrics

To evaluate our system, we use two approaches: 1) benchmarks using domain specific datasets, 2) the framework Ragas[3] which measure faithfulness, answer relevancy, and accuracy. It also allows to define specific adapted metrics to our needs. We also test and validate query syntax, accuracy of generated queries, and overall system performance. We envision to extend and automate a set of tests to be applied on the retrievers, the generator and the distributed approach vs monolithic agent.

VIII. Conclusion

Latest advances in Generative AI combined with Knowledge Engineering have made innovation possible to face unprecedented challenges in Contemporary archival practices

This work introduces an innovative approach to the archivistic application domain. It combines Advanced RAG and knowledge graphs techniques to maximize context and retrieve precise information to the user. The first step of the project consisted in setting up a conversational agent on an initial dataset (ICA) and the second step will involve doctoral research to add heterogenous data sources and to set up an advanced multi-agent architecture to tackle the distributed nature of the domain and provide the best way to explore archival use cases [11].

Acknowledgment

Special thanks to Aurèle Nicolet and ArchiLab team, Pierre-Yves Burgi from University of Geneva, Luciana Duranti and Muhammad Abdul-Mageed, InterPares Trust AI coleaders and our partners at ICA and PIAF. For their support and contributions. The authors would like to thank also Maryam Ablouâ for her work on the PIAF dataset.

References

[1] G. Colavizza, T. Blanke, C. Jeurgens, and J. Noordegraaf, "Archives and AI: an overview of current debates and future perspectives," *J. Comput. Cult. Herit.*, vol. 15, no. 1, pp. 1–15, Feb. 2022, doi: 10.1145/3479010.

[2] A. Hawkins, "Archives, linked data and the digital humanities: increasing access to digitised and born-digital archives via the semantic web," *Arch. Sci.*, vol. 22, no. 3, pp. 319–344, Sept. 2022, doi: 10.1007/s10502-021-09381-0.

[3] A. Marquet, "A maturity model for measuring digital transformation of archives and libraries," *Qualitative and Quantitative Methods in Libraries*, vol. 10, no. 3, pp. 269–282, Oct. 2021.

[4] J. Sheridan and C. Foster, "'digitalising a national archive': interview with john sheridan, digital director at the national archives, UK," *AI Soc.*, vol. 39, no. 2, pp. 665–668, Apr. 2024, doi: 10.1007/s00146-022-01510-2.

[5] I. Schellnack-Kelly and M. Modiba, "Developing smart archives in society 5.0: leveraging artificial intelligence for managing audiovisual archives in Africa," *Inf. Dev.*, vol. 41, no. 3, pp. 626–641, Sept. 2025, doi: 10.1177/02666669241286224.

[6] S. J. Russell and P. Norvig, Eds., *Artificial Intelligence: A Modern Approach*, 4th ed. 2022. [Online]. Available: https://aima.cs.berkeley.edu/global-index.html

[7] A. Z.Tsague, E. T. Fute, and L. P. Fotso, "Modeling of an environment for electronic archiving using a multi-agent approach," *Int. J. Appl. Inf. Syst.*, vol. 5, no. 7, pp. 32–37, May 2013, doi: 10.5120/ijais13-450941.

[8] J. Pellegrino, M. Maggiora, and W. Allasia, "A multi-agent approach for autonomous digital preservation," in *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, Turin, Italy: IEEE, June 2015, pp. 1–6. doi: 10.1109/ICMEW.2015.7169866.

[9] S. Aryal *et al.*, "Leveraging multi-AI agents for cross-domain knowledge discovery," Apr. 12, 2024, *arXiv*: arXiv:2404.08511. doi: 10.48550/arXiv.2404.08511.

[10] Portail International Archivistique Francophone, "Tous les cours," Portail International Archivistique Francophone. Accessed: Oct. 31, 2025. [Online]. Available: https://www.piaf-archives.org/tous-les-cours

[11] W. Ramli, *Ask_archilab tool*, (June 09, 2025). [Online Video]. Available: https://www.youtube.com/watch?v=NNoAd_HMDRM

[12] Multi-Agent GraphRAG: A Text-to-Cypher Framework for Labeled Property Graphs

[13] X. Comtesse et al. "Bots and Robots", october 2025, BOTS AND ROBOTS | Georg Editeur

---

Ragas[3]