

# Artificial Intelligence, LLM and RAG supporting Library and Information Science: challenges and opportunities

Giorgia Di Marcantonio

University of Macerata (Italy)

**Abstract.** ICTs are transforming the methods Library and Information Science uses to organize and disseminate knowledge. This transformation is evident in the adoption of new descriptive standards (RDA/RIC-CM) and numerous research projects (such as INTERPARES TRUST AI) that seek to assess and manage the impact of technologies, especially artificial intelligence, on archival science and librarianship within a global framework. The contribution offers insights into the future of knowledge description and cataloging, emphasizing the enhanced access to informational resources enabled by the use of LLM and RAG in digital environments designed for exploring cultural heritage

**Keywords.** Artificial Intelligence; Archives; Library and Information Science; LLMs; RAG;

The critical situation that led to the National Recovery and Resilience Plan (PNRR)<sup>1</sup> has secured substantial funding for Italy's "digital transition". The Plan dedicates significant financial resources to "Digitalization, innovation, competitiveness, culture, and tourism," amounting to €40.29 billion (Di Marcantonio and Valacchi 2022). This investment includes a major focus on cultural heritage, aiming to establish a vast repository of digitized resources to be housed in a National Digital Library (Cerullo and Negri 2023).

Researchers seeking information on Italy's cultural heritage online already have access to numerous digital platforms and information systems. These include the National Archival System (SAN)<sup>2</sup>, the Open Public Access Catalogue (OPAC-SBN)<sup>3</sup>, Internet Culturale<sup>4</sup>, Alfabetica<sup>5</sup>, the Information System of the State Archives<sup>6</sup>, Digital Archive<sup>7</sup>, Manus Online<sup>8</sup>, Cultura Italia<sup>9</sup>, and the General Catalog of Cultural Heritage<sup>10</sup>, among others. Each platform has its own navigation paths, and while some offer direct access to digitized cultural resources, others provide only summary information, indicating whether a resource exists, where it is stored, and how it can be accessed within cultural institutions. In Italy, there is a lack of comprehensive studies on user satisfaction with these platforms (Feliciati 2016). The few studies that do exist provide limited insight, leaving a significant gap in understanding how effectively these resources serve their users (Craig 1998; Chapman 2010). Navigating these platforms can be especially difficult for non-expert users due to the prevalent use of technical language that can be hard to understand for those not well-versed in the field. This complexity not only restricts access to valuable cultural information, but also deters wider public engagement. Simplifying user interfaces and providing more intuitive navigation and search capabilities could greatly improve the user experience and make these resources more accessible (Djamasbi, Siegel, and Tullis 2010).

Digital platforms frequently lack the cultural mediation between scholars and resources that is typically present in archives, libraries, and museums (Michetti 2020). This highlights the necessity of thoughtfully presenting and organizing cultural knowledge online. The target group of users of these platforms is varied, encompassing students, scholars, researchers, and ordinary citizens who deserve easy access to their cultural heritage. However, creating navigation paths that satisfy all potential users of digital cultural resources is a significant challenge. The resources are varied, different description standards are employed, and each scientific field has its unique requirements that it is reluctant to compromise. Enhancing access systems could greatly benefit from the use of Artificial Intelligence (AI). Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) could facilitate natural interactions between users and platforms (Kasneci et al. 2023).

Large Language Models (LLMs) are a type of AI that can understand and generate human language. Trained on vast amounts of text data, they can perform various language tasks, such as translation, summarization, and question-answering. Models like OpenAI's GPT-4 can understand context, infer meaning, and generate coherent and contextually relevant responses, making them powerful tools for enhancing user interaction with digital platforms.

The integration of Artificial Intelligence (AI) in managing cultural resources presents both exciting opportunities and significant challenges. AI systems, especially Large Language Models (LLMs), have the potential to produce inaccurate results, known as hallucinations, which can stem from biases in the training data (Ji et al. 2023). Additionally, AI systems can perpetuate stereotypes and manipulate information, raising ethical concerns. In the context of cultural resource management, these limitations are particularly problematic because the goal is to provide reliable, scientifically validated information for research and self-education.

One major issue with LLMs is their tendency to generate plausible-sounding but incorrect or misleading information. This is especially critical when dealing with cultural heritage, where accuracy and reliability are paramount. Users need to trust that the information they access about cultural artifacts, historical documents, or scholarly works is both accurate and authoritative.

To mitigate the issues inherent in LLMs, a promising approach is combining them with Retrieval-Augmented Generation (RAG) systems (Lewis et al. 2020). RAG systems enhance the relevance and accuracy of generated responses by integrating a retrieval mechanism that draws upon a pre-defined knowledge base. This dual approach ensures that responses are grounded in verifiable data.

In practical terms, a RAG system starts by retrieving relevant documents or data from a curated database or a selected set of resources based on the user's query. This retrieved information then guides the text generation process, ensuring that the responses are contextually appropriate and factually accurate.

A significant advantage of RAG systems is their ability to address and mitigate biases inherent in training data. By relying on carefully selected databases, RAG systems ensure

the information used to generate accurate and free from distortions responses. This is crucial in cultural resource management, where information integrity is paramount.

To further enhance reliability, these selected databases are often converted into a vectorized format. This process involves converting texts into numerical sequences using embedding algorithms, enabling the system to perform comparisons based on vector similarity. By comparing the input vector with those representing information in the database, the system can identify the most relevant resources to answer the user's query. Techniques such as cosine similarity are commonly used to ensure the most pertinent information is retrieved and utilized.

As AI continues to be integrated into cultural resource management, it is essential to address the ethical implications. The capacity of AI to influence user opinions through language manipulation is a double-edged sword. While it can make interactions more engaging and informative, it also has the potential to mislead if not carefully managed (Chen et al. 2024). Ensuring transparency in AI operations and maintaining rigorous standards for data accuracy and bias mitigation are critical for fostering trust and reliability.

Continuous user feedback and systematic studies on user satisfaction and accessibility are necessary to refine these digital platforms. Understanding the diverse needs of users, from scholars to everyday citizens, will aid in designing more intuitive and inclusive interfaces, making cultural heritage more accessible to all.

LLMs, whether used alone or with RAG integration, do not overhaul the existing systems used for managing cultural resources. Catalog entries, archival descriptions, and inventories should—and must—remain accessible to users. What these technologies enhance are the points of access to information, facilitating that initial exploration or navigation when it's unclear what resources are available on a given topic. RAG systems increase the likelihood of serendipitous discoveries and provide more structured responses to user queries.

Incorporating these technologies into knowledge management environments can significantly reduce the time needed to find and utilize resources. However, they should be viewed as supplementary features rather than replacements for the existing systems. By adding these advanced functionalities, users can benefit from faster, more intuitive access to the wealth of information contained within traditional cataloging and archival systems, without losing the depth and reliability these systems provide (Di Marcantonio 2024).

Integrating LLMs and RAG systems could greatly enhance access to cultural resources on digital platforms. However, it is crucial to invest in the training of industry professionals, data modeling, and the technological tools available to cultural institutions. The digital transformation requires significant investments, which have been made possible in Italy through the PNRR. However, embracing digital methodologies is equally crucial. To simply have the tools is insufficient; a "digital mindset" is essential to produce user-friendly and accurate contents that are potentially available in long-term for future generations—the primary target of the NEXGenerationEU funds.

## References

Cerullo, Luigi, e Antonella Negri. 2023. «L'infrastruttura software per il patrimonio culturale (ISPC) come abilitatore di un Ecosistema digitale nazionale del patrimonio culturale». *Digitalia* 18 (1): 38–50. <https://doi.org/10.36181/digitalia-00059>.

Chapman, Joyce Celeste. 2010. «Observing Users: An Empirical Analysis of User Interaction with Online Finding Aids». *Journal of Archival Organization* 8 (1): 4–30. <https://doi.org/10.1080/15332748.2010.484361>.

Chen, Xiaoyang, Ben He, Hongyu Lin, Xianpei Han, Tianshu Wang, Boxi Cao, Le Sun, e Yingfei Sun. 2024. «Spiral of Silences: How is Large Language Model Killing Information Retrieval? A Case Study on Open Domain Question Answering». <https://doi.org/10.48550/ARXIV.2404.10496>.

Craig, Barbara Lazenby. 1998. «Old Myths in New Clothes: Expectations of Archives Users». *Archivaria* 45 (February): 118–26. <https://archivaria.ca/index.php/archivaria/article/view/12228>.

Di Marcantonio, Giorgia. 2024. «Artificial Intelligence, Large Language Models (LLMs), and Retrieval-Augmented Generation (RAG). New Tools for Accessing Archival and Bibliographic Resources». *Bibliothecae.it* 13 (1).

Di Marcantonio, Giorgia, e Federico Valacchi. 2022. «Tra dematerializzazione e next generation: i sistemi documentari nel PNRR». *Rivista elettronica di diritto, economia, management* 2:16–32.

[https://www.clioedu.it/documenti/eventi-live-ondemand/rivista-elettronica/Rivista-elettronica-2\\_2022.pdf](https://www.clioedu.it/documenti/eventi-live-ondemand/rivista-elettronica/Rivista-elettronica-2_2022.pdf)

Djamasbi, Soussan, Marisa Siegel, e Tom Tullis. 2010. «Generation Y, Web Design, and Eye Tracking». *International Journal of Human-Computer Studies* 68 (5): 307–23. <https://doi.org/10.1016/j.ijhcs.2009.12.006>.

Feliciati, Pierluigi. 2016. «L'usabilità degli ambienti bibliotecari e archivistici digitali come requisito di qualità: contesto, modelli e strumenti di valutazione». *JLIS.it* 7 (1): 113–30. <https://doi.org/10.4403/jlis.it-11512>

Ji, Ziwei, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, e Pascale Fung. 2023. «Survey of Hallucination in Natural Language Generation». *ACM Computing Surveys* 55 (12): 1–38. <https://doi.org/10.1145/3571730>.

Kasneci, Enkelejda, Kathrin Sessler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, et al. 2023. «ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education». *Learning and Individual Differen-*

ces 103. <https://doi.org/10.1016/j.lindif.2023.102274>.

Lewis, Patrick, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, et al. 2020. «Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks». [Preprint] <https://doi.org/10.48550/ARXIV.2005.11401>.

Michetti, Giovanni. 2020. «Il mondo come puzzle. I beni culturali nel web». *Digitalia Web Rivista del digitale nei beni culturali* 15 (1): 26–42. <https://doi.org/10.36181/digitalia-00002>

## Notes

<sup>1</sup><https://www.governo.it/sites/governo.it/files/PNRR.pdf> (Last access: 6/24/24).

<sup>2</sup><http://san.beniculturali.it> (Last access: 6/24/24).

<sup>3</sup><https://opac.sbn.it> (Last access: 6/24/24).

<sup>4</sup><https://www.internetculturale.it/> (Last access: 6/24/24).

<sup>5</sup><https://alphanumerica.it/web/alphanumerica/> (Last access: 6/24/24).

<sup>6</sup><https://sias.archivi.beniculturali.it/> (Last access: 6/24/24).

<sup>7</sup><https://www.archiviodigitale.icar.beniculturali.it> (Last access: 6/24/24).

<sup>8</sup><https://manus.iccu.sbn.it/> (Last access: 6/24/24).

<sup>9</sup><https://www.culturaitalia.it/> (Last access: 6/24/24).

<sup>10</sup><https://catalogo.beniculturali.it> (Last access: 6/24/24).



## Author

Giorgia Di Marcantonio [g.dimarcantonio@unimc.it](mailto:g.dimarcantonio@unimc.it)

Giorgia Di Marcantonio is a researcher in Archival and Library Science, (SSD M-5T0/08) at the University of Macerata. Her main research areas explore the relationship between archives and ICTs, with a special focus on international standards and models for information retrieval. She is a member of the scientific committee for the Italian Journal of Library, Archives and Information Science (JLIS.it) and has published numerous articles in sector-specific scientific journals. Recently, she has been involved with the international project InterPARES Trust AI (2021-2026), which aims to delineate the risks and benefits associated with the use of artificial intelligence in record management.