

Retrieval-Augmented Generation for Exploring the Catalogues of Pistoia Museums

Luca Bindini^{1,*}, Marco Lippi¹ and Simone Marinai¹

¹AI Lab, Department of Information Engineering, University of Florence, via di Santa Marta 3, 50139 Florence, Italy

Abstract

This paper presents a research project developed in collaboration with a network of museums in the Province of Pistoia. The project aims to design and implement a chatbot that enables citizens and researchers to explore museum collections, search for cultural objects, and plan thematic visits. The proposed system combines an open-weight Large Language Model with a Retrieval-Augmented Generation architecture whose knowledge base is constructed from museum catalogue records.

Keywords

Large Language Models, Retrieval Augmented Generation, Catalogue Exploration, Museum Collections

1. Introduction

During the last decade, there has been a growing emphasis on designing and developing new tools and systems, based on Artificial Intelligence, to improve user experience in visiting museums and heritage collections [1, 2]. Within this context, Large Language Models (LLMs), including multimodal models, clearly represent a key instrument, for their capability of handling large and heterogeneous data collections, and for allowing interaction with end-users in plain natural language [3]. In fact, one of the major difficulties in developing applications for the world of Cultural Heritage is that data collections are very often unstructured, fragmented in different local databases and management systems, sometimes even in libraries or archives that are not yet fully digitalized [4]. From this point of view, significant effort is required in order to make these collections homogeneous, standardized, and fully accessible.

In this paper, we present a project developed by the Artificial Intelligence Laboratory of the University of Florence in collaboration with Sistema Museale Pistoiese (SIMUP), a network of museums in the Province of Pistoia, in central Italy. The project, named AIACE (Artificial Intelligence for Advanced Catalogue Exploration of Pistoia Museums) aims to develop a chatbot for citizens and researchers, for the exploration of the catalogues of all the museums belonging to the consortium, thus offering a single entry point for tourists, as well as domain experts, to discover the different collections, finding useful connections and links to plan a visit, prepare an itinerary, or to search for specific details of exhibits and artworks. The proposed solution, still under development, is based on a Retrieval-Augmented Generation (RAG) architecture coupled with an open-weight LLM, indexing the currently acquired catalogue records locally to improve access to the different collections. The contribution lies in the real multi-institutional setting and the groundwork for progressive data integration.

In the remainder of the paper, we first describe the SIMUP network of museums (Section 2); then, we illustrate the heterogeneous nature of the available data (Section 3) and the designed system architecture (Section 4); finally, we conclude the paper with a discussion of research challenges and opportunities (Section 5).

4th Italian Workshop on Artificial Intelligence for Cultural Heritage (IAI4CH 2026, <https://ai4ch.di.unito.it/>), co-located with the 24th International Conference of the Italian Association for Artificial Intelligence (AIIA 2026). 6–9 October 2026, Perugia, Italy.

*Corresponding author.

✉ luca.bindini@unifi.it (L. Bindini); marco.lippi@unifi.it (M. Lippi); simone.marinai@unifi.it (S. Marinai)

ORCID 0009-0008-7024-2888 (L. Bindini); 0000-0002-9663-1071 (M. Lippi); 0000-0002-6702-2277 (S. Marinai)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

2. The SIMUP Museum Network

The Sistema Museale Pistoiese (SIMUP) is a territorial network comprising thirteen public and private museums located in the municipalities of Pistoia, Pescia, and Monsummano Terme, in Tuscany, Italy.¹ Rather than representing a single collection, SIMUP connects institutions with different administrative structures, histories, audiences, and cataloguing practices. Taken together, the museums document the cultural development of the Pistoia area from archaeology and medieval art to modern and contemporary artistic production, as well as the history of medicine, local manufacturing, and the transformation of the territory. Connecting these collections requires handling differences in terminology and descriptive detail.

Museums in Pistoia. *Museo Civico d'Arte Antica in Palazzo Comunale.* Located in the fourteenth-century Palazzo degli Anziani, it is the oldest and largest civic museum in Pistoia. Its collection comprises more than three hundred paintings, sculptures, and applied-art objects documenting the artistic history of the city from the twelfth to the nineteenth century. Particularly significant groups include medieval gold-ground paintings, Renaissance altarpieces, and works from the Puccini collection.

Museo dell'Antico Palazzo dei Vescovi. The museum is housed in a building documented since the eleventh century and closely connected with the religious and civic history of Pistoia. Its collections include archaeological evidence of the urban stratification of the city, liturgical furnishings from the Cathedral, the sixteenth-century Millefiori tapestry, the seventeenth-century paintings from the Bigongiari collection, and wall paintings by Giovanni Boldini.

Palazzo de' Rossi – Collezioni del Novecento. The permanent exhibition presents a selection of twentieth- and twenty-first-century art connected with Pistoia. Works by artists such as Galileo Chini, Marino Marini, Fernando Melani, Gualtiero Nativi, Mario Nigro, and Umberto Boccioni document the development of local artistic research in dialogue with broader Italian movements.

Museo di San Salvatore. Created inside the restored church of San Salvatore, documented since the tenth century, the museum combines archaeology, architecture, and medieval art. The visit includes the remains discovered during the excavations, a medieval burial area, and a late-thirteenth-century fresco representing the Lamentation over the Dead Christ. Multimedia and tactile resources provide different levels of interpretation for visitors.

Palazzo Buontalenti. This late-sixteenth-century building is used primarily for temporary exhibitions organized by Pistoia Musei. Its programme focuses especially on modern and contemporary art, photography, and transversal cultural themes, making it a flexible exhibition venue rather than a museum defined by a single permanent collection.

Museo dello Spedale del Ceppo. Located in the former hospital complex, the museum documents more than seven centuries of medical and charitable activity in Pistoia. Its itinerary includes the history of the hospital, a collection of surgical instruments, the eighteenth-century anatomical theatre, and an interpretation of the Renaissance glazed-terracotta frieze decorating the external loggia.

Museo del Novecento e del Contemporaneo di Palazzo Fabroni. The museum houses the permanent modern and contemporary art collection of the Municipality of Pistoia. Its chronological itinerary extends from the 1920s to the present and includes monographic sections devoted to artists such as Mario Nigro, Fernando Melani, Gualtiero Nativi, and Agenore Fabbri, together with works related to Arte Povera, Conceptual Art, Minimal Art, and Visual Poetry.

Casa-studio Fernando Melani. The house-studio of Fernando Melani preserves more than 2,800 works and experiments in the spatial arrangement established by the artist. Objects, metals, wires, surfaces, and everyday materials document Melani's research and his distinctive engagement with the principal experimental movements of the second half of the twentieth century.

Museums in Pescia. *Gipsoteca Libero Andreotti.* Located in the medieval Palazzo del Podestà, the Gipsoteca is dedicated to the sculptor Libero Andreotti (1875–1933). Its collection contains approximately

¹<https://www.sistemamusealepistoiese.it/musei>

250 sketches, plaster casts, models, and fragments, together with the artist's archive of more than 5,000 documentary units. These materials make it possible to reconstruct both Andreotti's creative process and his relationships with the Italian and European artistic environment of the early twentieth century.

Museo Civico di Palazzo Galeotti. The museum originated at the end of the nineteenth century from the collection of artworks gathered from churches, public buildings, and private donations in Pescia. Its holdings include paintings, sculptures, drawings, decorative objects, and archaeological materials, with particular attention to works connected with the city and its territory.

Museo della Carta di Pescia. The museum is housed in the former Le Carte paper mill and preserves the history of handmade paper production in the Pescia area, documented from the end of the fifteenth century. Its collections include paper moulds, watermark waxes, punches, metal sheets, stamps, machinery, and the historical archive of the Magnani paper mills, providing an important record of local industrial and manufacturing history.

Museums in Monsummano Terme. *Museo di Arte Contemporanea e del Novecento.* Located in Villa Renatico Martini, the museum preserves the civic collection *Il Renatico*, consisting of paintings, sculptures, and graphic works from the twentieth and twenty-first centuries. The collection documents Tuscan artistic production while also including artists active at the national and international levels.

Museo della Città e del Territorio. Housed in the former Osteria dei Pellegrini, the museum presents the historical development of Monsummano Terme and the Valdinievole. Its interdisciplinary itinerary includes geology, palaeontology, archaeology, the Padule di Fucecchio, astronomical and topographical instruments, traditional manufacturing, thermal activities, migration, and the memory of the Padule di Fucecchio massacre. A notable exhibit is also the Fontenuova Treasury, a collection of votive offerings and sacred furnishings associated with the Sanctuary of Maria Santissima della Fontenuova.

This variety creates opportunities for cross-collection exploration, but also challenges for data normalization, semantic retrieval, and coherent thematic narratives. The initial AIACE scope includes ten of the thirteen SIMUP museums, listed in Table 1. Museo dello Spedale del Ceppo, Museo del Novecento e del Contemporaneo di Palazzo Fabroni, and Casa-studio Fernando Melani belong to SIMUP but are excluded from the current data-integration phase. Thus, the Museo Civico d'Arte Antica in Palazzo Comunale is the only civic museum in Pistoia included at this stage.

3. Data Availability

An initial survey was conducted to assess the availability and the structure of the catalogue resources associated with the ten museums included in the initial AIACE scope. The survey revealed substantial differences among institutions, even though with some common ground. Some collections, in fact, are represented in the Italian General Catalogue of Cultural Heritage,² managed by the Ministry of Culture. Other items are instead available in a variety of different formats: institutional websites, internal databases and spreadsheets, or even printed catalogues.

Table 1 summarizes the museums currently considered by AIACE: it distinguishes between records available through the ministerial catalogue and additional resources that have been identified but are not yet integrated into the AIACE knowledge base.

The current corpus comprises 3,511 ministerial records associated with five of the ten museums listed in Table 1. These records constitute the first subset that has been systematically acquired and normalized. The remaining catalogue resources reported in the last column describe the planned expansion of the knowledge base and are not yet used by the current prototype. Some of these additional sources may partially overlap with records already available through the ministerial catalogue. Their integration will therefore require record linkage, deduplication, and museum-specific normalization before the actual number of additional cultural objects can be determined. During the preprocessing stage, we map each catalogue record to a common structured representation containing the information used for retrieval

²<https://catalogo.beniculturali.it/>

Table 1

Availability of catalogue data for the ten museums included in the initial AIACE scope. The second column reports records currently available through the Italian General Catalogue of Cultural Heritage. The third column describes additional catalogue resources identified but not yet integrated into AIACE.

Museum	Ministerial records	Additional resources
<i>Pistoia</i>		
Museo Civico d'Arte Antica in Palazzo Comunale	217	Non-public spreadsheet and additional descriptive records
Museo dell'Antico Palazzo dei Vescovi	131	Non-public spreadsheet
Palazzo de' Rossi – Collezioni del Novecento	–	Non-public spreadsheet
Museo di San Salvatore	–	Non-public spreadsheet
Palazzo Buontalenti	–	Non-public spreadsheet
<i>Pescia</i>		
Gipsoteca Libero Andreotti	352	Additional descriptive records
Museo Civico di Palazzo Galeotti	2,406	Additional descriptive records
Museo della Carta di Pescia	405	Image records available through Europeana, partial catalogue on museum website, Magnani Archive in a private database
<i>Monsummano Terme</i>		
Museo di Arte Contemporanea e del Novecento	–	Complete catalogue available on the museum website
Museo della Città e del Territorio	–	Partial online resources and a complete printed catalogue
Current AIACE corpus	3,511	To be integrated

and generation. Specifically, we retain the title of the object, author or attribution, dating, object type, subject, materials and techniques, measurements, museum, location, and historical or critical notes. When this information is unavailable in the source record, the corresponding fields are left empty rather than being automatically inferred. Accordingly, each catalogue record is treated as an independent document in our system, and its structured metadata are converted into a textual description such as:

Title: Popolana che saluta
 Author: Libero Andreotti
 Dating: 1919-1920
 Object type: Sculpture
 Materials and techniques: Plaster, modelling
 Museum: Gipsoteca Libero Andreotti
 Historical and critical notes: ...

This representation preserves the semantic unity of each catalogue entry without additional chunking. Museum and location fields support institutionally coherent itineraries, although practical feasibility still requires validation.

The five AIACE museums not represented in the current corpus are associated with heterogeneous resources whose availability and level of detail vary substantially: some data require institutional authorization, some are available only through museum-specific websites or shared spreadsheets, and others remain accessible primarily through printed catalogues. Their integration will require dedicated extraction and normalization procedures, as well as validation with institutions and domain experts.

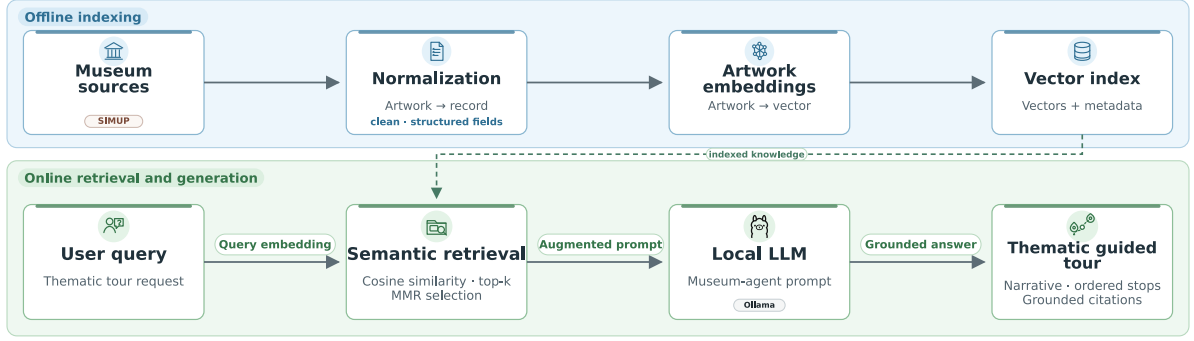


Figure 1: Overview of the proposed RAG-based system architecture. Top (light blue): museum records are normalized and indexed during the offline phase. Bottom (light green): at query time, semantic retrieval followed by Maximal Marginal Relevance selects relevant and diverse artworks, which are provided to a local large language model to generate a grounded thematic guided tour.

4. System architecture

Figure 1 provides an overview of the proposed architecture. The system is organized into an offline phase, in which museum records are collected, pre-processed, normalized, and indexed, and an online phase, in which users can provide natural-language requests that are elaborated to retrieve and select relevant artworks before generating a thematic guided tour.

Each artwork is treated as a single document and represented using the normalized catalogue fields described in Section 3. Since museum records are generally short and refer to a single cultural object, no additional chunking strategy is applied. During indexing, *qwen3-embedding* embeds each artwork description. Retrieval uses cosine similarity; no external vector database is required. Indexing is repeated when the corpus or embedding model changes. Both models are accessed through Ollama,³ and the local generative model is *qwen3.5*.

At query time, the user request is encoded using the same embedding model. Cosine similarity is then used to retrieve an initial set of candidate artworks. A retrieval strategy based exclusively on similarity may return several artworks with nearly identical subjects, materials, or catalogue descriptions: although these records may all be relevant to the query, their redundancy can reduce the quality of a thematic itinerary. Hence, the system applies Maximal Marginal Relevance (MMR) [5] to balance relevance to the user request with diversity among the selected artworks. Let C be the initial set of candidate artworks retrieved through cosine similarity, and let S be the set of artworks already selected for the final context. At each iteration, the next artwork d^* is selected as

$$d^* = \arg \max_{d \in C \setminus S} \left[\lambda \text{sim}(q, d) - (1 - \lambda) \max_{s \in S} \text{sim}(d, s) \right]. \quad (1)$$

Here, q denotes the user query and $\text{sim}(\cdot, \cdot)$ is the cosine similarity between embedding vectors. The first term rewards artworks that are relevant to the requested theme, whereas the second term penalizes candidates that are too similar to artworks already included in S . During the first iteration, when S is empty, the redundancy term is set to zero. The parameter $\lambda \in [0, 1]$ controls the trade-off between relevance and diversity. Values close to one primarily favour similarity to the user query, whereas lower values encourage a more heterogeneous selection. In thematic tour generation, MMR therefore reduces the retrieval of nearly equivalent objects while preserving coherence with the requested theme.

The final records are assigned temporary identifiers and included in the augmented prompt. The generative model is instructed to act as an expert in art history, museology, cultural mediation, and guided-tour design. It produces a direct response to the user, organizing the selected artworks into a coherent thematic narrative and associating factual statements with the corresponding identifiers.

³<https://ollama.com/>

Qualitative example. For “*Dove posso trovare opere in bronzo?*” (“Where can I find bronze artworks?”), the response presents different records from three museum venues in Pescia and Pistoia. At *Museo Civico di Palazzo Galeotti* in Pescia, it recommends bronze objects including *Madonna*, a bust, and *putto*, a small sculpture. At *Gipsoteca Libero Andreotti*, also in Pescia, it includes *Brandano pescatore*, a plaster bust, while mentioning bronze versions of the same work. At *Museo Civico d’Arte Antica* in Pistoia, it proposes *medagliere*, an object combining bronze and wood. The response also mentions Carlini’s *vasca di bronzo con base*, a drawing depicting a bronze basin, preserved at Palazzo Galeotti. This example illustrates how the system connects records from different museums through a shared theme.

5. Discussion

This paper presents the AIACE project, its current data resources, and the designed RAG architecture. Since the system is still under development, a central part of future work will concern the definition of an appropriate validation methodology. The evaluation should consider both retrieval and generation, and it will necessarily involve human experts. In fact, retrieval performance can be assessed by measuring the relevance and diversity of the selected artworks, whereas generated answers and itineraries should be evaluated in terms of factual correctness, grounding in catalogue records, thematic coherence, and practical feasibility. These criteria will require the construction of a dedicated benchmark and the participation of museum professionals, whose assessments will be essential to determine whether the generated contents and itineraries are meaningful and useful to visitors.

Acknowledgments

The AIACE (Artificial Intelligence for Advanced Catalogue Exploration of Pistoia Museums) project was funded by Fondazione Cassa di Risparmio di Pistoia (Fondazione CariPT). We also thank all the SIMUP museums for their valuable contribution to the development of the project.

Declaration on Generative AI

During the preparation of this work, the author(s) occasionally used Gemini and ChatGPT for the following tasks: paraphrasing and rewording.

References

- [1] R. Palumbo, Enhancing museums’ attractiveness through digitization: An investigation of Italian medium and large-sized museums and cultural institutions, *International Journal of Tourism Research* 24 (2022) 202–215.
- [2] N. Raimo, I. De Turi, A. Ricciardelli, F. Vitolla, Digitalization in the cultural industry: evidence from Italian museums, *International Journal of Entrepreneurial Behavior & Research* 28 (2022) 1962–1974.
- [3] M. Reusens, A. Adams, B. Baesens, Large language models to make museum archive collections more accessible, *AI & SOCIETY* 40 (2025) 4485–4497.
- [4] B. Ranjgar, A. Sadeghi-Niaraki, M. Shakeri, F. Rahimi, S.-M. Choi, Cultural heritage information retrieval: past, present, and future trends, *IEEE Access* 12 (2024) 42992–43026.
- [5] J. Carbonell, J. Goldstein, The use of MMR, diversity-based reranking for reordering documents and producing summaries, in: *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, 1998, pp. 335–336.