From Stone to Signal: Heritage Lens

A Multimodal AI System for Archaeological Interpretation

Cultural Heritage Management and Conservation Conference
Larissa Terranova | Centro Studi Americanistici "Circolo Amerindiano" (CSACA)

Introduction

This project introduces a multimodal Retrieval-Augmented Generation (RAG) system designed to enhance interpretive research in digital archaeology. Heritage Lens enables natural language queries across archaeological texts and images. Developed as a proof-of-concept using *The Olmec Town* thesis on Mesoamerican urban settlements (1600–400 BCE), the prototype combines text embeddings, vision-language models, and generative AI to provide semantically grounded responses from archaeological sources. It offers a new interface for scholarly engagement with complex archaeological materials whilst maintaining academic attribution standards. Developed within the *Ruta de la Obsidiana* research initiative, Heritage Lens addresses both technical innovation and ethical considerations for AI applications in cultural heritage contexts. Moreover, the project engages with ethical questions around AI's use in indigenous cultural contexts and proposes respectful data handling practices for digital heritage applications.

Essential Terminology

📌 **CLIP** — Model linking text and images in shared space, enabling cross-modal similarity matching.

📌 **Cross-modal Retrieval** — Search across different media types, using text queries to find images or vice versa.

📌 **Groundedness** — Degree to which AI responses can be traced back to retrieved source materials.

📌 **LLM-as-a-Judge** — Using another language model to **evaluate the quality** of an AI's answer (checking accuracy, relevance, groundedness).

📌 **Multimodal AI** — AI processing multiple data types (text, images) simultaneously rather than single formats.

📌 **Olmec Civilisation** — Early Mesoamerican culture (c. 1600–400 BCE), foundational to subsequent developments.

📌 **RAG (Retrieval-Augmented Generation)** — AI retrieves from sources before generating answers, ensuring groundedness in original materials.

📌 **Semantic Search** — Information retrieval based on meaning rather than exact keyword matching.

Presentation Outline

Opening & Demonstration

- Conversational access to archaeological knowledge through AI
- Proof-of-concept prototype from Johns Hopkins Applied Generative AI course
- Video demo showing natural language queries with integrated text-image responses
- Academic context: *Ruta de la Obsidiana* international research project

The Challenge & Vision

- Problem: Heritage knowledge is fragmented across texts, maps, and images
- Current reality: Hours/days to locate specific details in dense documents
- Solution: Ask questions naturally, receive integrated answers instantly

How Heritage Lens Works

- **Three components:** Research Assistant (semantic text search), Visual Curator (image-text matching), Quality Checker (source attribution)
- **Simple workflow:** Search → Retrieve → Explain

Current Scope & Limitations

- Single-document prototype (Olmec thesis only)
- Multilingual queries supported
- Information access tool, not interpretation replacement
- Research assistant complementing traditional methods

Impact & Future Vision

- Democratises heritage knowledge for students, researchers, museum professionals
- From one thesis to museums, archives, classrooms worldwide
- Ethical AI promoting respectful cultural heritage engagement
- Bridge connecting past-present, research-public, words-images

---

Contact Information:
Larissa Terranova | l.terranova3@gmail.com
QR code available for feedback and LinkedIn contact

LinkedIn

Feedback Form