AI & ANTIQUITY

Editey by

CIAW | CENTER FOR INNOVATION IN ANCIENT WORLDS

With support from

UAB
Universitat Autònoma de Barcelona

AI & Antiquity
Journal of Teaching and Technology in Ancient Studies

Volume 2 (2026) - Issue 1

Edited by the *Center for Innovation in Ancient Worlds*

ABOUT THE JOURNAL

When References Mislead: Verification, AI Attribution, and Academic Bullying in Scholarly Evaluation

Carlos Heredia Chimeno

Editor-in-Chief
Universitat Autònoma de Barcelona
carlos.heredia@uab.cat
0000-0003-2866-5883

1. A SILENT PROBLEM IN PLAIN SIGHT

If the first volume of *AI & Antiquity* was born from the need to rethink how we teach and learn in the age of artificial intelligence, this second volume arises from a more uncomfortable realisation: AI does not only challenge assessment or authorship; it challenges the very infrastructure of scholarly trust, and the centrality of source verification.

Among the many issues raised by the widespread use of generative AI in academic contexts, one has proven particularly insidious: the creation of bibliographic references that do not exist. These are not necessarily fabricated with malicious intent, nor always introduced by students attempting to deceive. More often, they are generated silently, plausibly, and convincingly by systems trained to reproduce the surface features of scholarly discourse: titles that sound real, journals that almost exist, DOIs that follow correct patterns but lead nowhere, and authors whose names resonate with the field but have never written the cited work.

What makes this phenomenon especially troubling is not its novelty, but its subtlety. Unlike plagiarism, which leaves detectable traces, or factual errors that may be identified through disciplinary knowledge, false bibliography often passes unnoticed precisely because it looks correct (see Walters and Wilder, 2023; Liu, D'Elia and Palermo, this issue). Worse still, it frequently emerges in contexts perceived as low-risk: when asking AI to "format references", "complete missing citations", or "standardise a bibliography according to a style guide". These are tasks that many of us, as researchers, reviewers, or editors, have already delegated at least once, often assuming that the operation is merely cosmetic. This, ultimately, is the crux of the problem: bibliographic hallucination is not merely a student issue. It is a structural risk that affects the entire academic ecosystem, precisely because it exploits the conventions on which scholarship relies.

CARLOS HEREDIA CHIMENO

Actually, generative AI models are not databases. They do not "know" whether a reference exists; they predict text. When prompted to generate or adjust bibliographic entries, they do exactly what they are designed to do: produce something that looks right. The danger lies in our tendency to confuse formal correctness with epistemic reliability, as if the mere presence of a well-formed citation were synonymous with traceability and truth.

In the Humanities, and particularly in Ancient Studies, bibliography is not an accessory. It is the backbone of scholarly argumentation. References are not decorative footnotes; they are claims of verification—signals that knowledge can be traced, contested, and reassessed. When that chain is broken, even unintentionally, credibility collapses, and the reader is deprived of the very mechanism that makes scholarship cumulative rather than rhetorical. What is new in the AI context is the scale and speed at which such errors can propagate. A single fabricated reference, once cited, reformatted, exported to reference managers, or reused in teaching materials, can circulate widely before anyone notices the absence of an original source. This dynamic is amplified by the growing reliance on automated workflows, citation generators, and AI-assisted writing environments that prioritise efficiency over verification. The result is a paradox that we can no longer ignore: tools designed to assist scholarship can, if used uncritically, undermine the very standards they are meant to support.

2. VERIFICATION, DUE PROCESS, AND AI-RELATED ACADEMIC BULLYING

The response to this situation cannot be a return to technological rejection, nor can it rely solely on warnings directed at students. Instead, it requires a collective recalibration of scholarly habits—one that recognises that, in an AI-mediated environment, verification must be treated not as an optional final step but as the core scholarly practice linking responsibility, traceability, and human accountability. In this respect, few approaches are more instructive than the perspective advanced by Solga and Sarwar (2026, this issue) whose emphasis on procedural transparency and evidentiary anchoring foregrounds verification as a constitutive scholarly act rather than a post *hoc* corrective.

First, we must acknowledge that bibliographic vigilance is now a core scholarly skill, on par with source criticism or methodological transparency. Checking references is no longer a cosmetic final operation; it is the epistemic labour that distinguishes human scholarship from automated plausibility. Second, we must resist the temptation to treat AI as a neutral technical assistant. Delegating tasks such as citation completion, bibliography generation, or reference formatting without verification is not a time-saving shortcut; it is a transfer of responsibility to a system that cannot bear it. The historian, archaeologist, or philologist remains accountable for every reference that appears under their name, just as they remain accountable for the evidence, reasoning, and interpretative claims constructed from

5

© AI & Antiquity. Volume 2, Issue 1, 2026.

CARLOS HEREDIA CHIMENO

Actually, generative AI models are not databases. They do not "know" whether a reference exists; they predict text. When prompted to generate or adjust bibliographic entries, they do exactly what they are designed to do: produce something that looks right. The danger lies in our tendency to confuse formal correctness with epistemic reliability, as if the mere presence of a well-formed citation were synonymous with traceability and truth.

In the Humanities, and particularly in Ancient Studies, bibliography is not an accessory. It is the backbone of scholarly argumentation. References are not decorative footnotes; they are claims of verification—signals that knowledge can be traced, contested, and reassessed. When that chain is broken, even unintentionally, credibility collapses, and the reader is deprived of the very mechanism that makes scholarship cumulative rather than rhetorical. What is new in the AI context is the scale and speed at which such errors can propagate. A single fabricated reference, once cited, reformatted, exported to reference managers, or reused in teaching materials, can circulate widely before anyone notices the absence of an original source. This dynamic is amplified by the growing reliance on automated workflows, citation generators, and AI-assisted writing environments that prioritise efficiency over verification. The result is a paradox that we can no longer ignore: tools designed to assist scholarship can, if used uncritically, undermine the very standards they are meant to support.

2. VERIFICATION, DUE PROCESS, AND AI-RELATED ACADEMIC BULLYING

The response to this situation cannot be a return to technological rejection, nor can it rely solely on warnings directed at students. Instead, it requires a collective recalibration of scholarly habits—one that recognises that, in an AI-mediated environment, verification must be treated not as an optional final step but as the core scholarly practice linking responsibility, traceability, and human accountability. In this respect, few approaches are more instructive than the perspective advanced by Solga and Sarwar (2026, this issue) whose emphasis on procedural transparency and evidentiary anchoring foregrounds verification as a constitutive scholarly act rather than a post *hoc* corrective.

First, we must acknowledge that bibliographic vigilance is now a core scholarly skill, on par with source criticism or methodological transparency. Checking references is no longer a cosmetic final operation; it is the epistemic labour that distinguishes human scholarship from automated plausibility. Second, we must resist the temptation to treat AI as a neutral technical assistant. Delegating tasks such as citation completion, bibliography generation, or reference formatting without verification is not a time-saving shortcut; it is a transfer of responsibility to a system that cannot bear it. The historian, archaeologist, or philologist remains accountable for every reference that appears under their name, just as they remain accountable for the evidence, reasoning, and interpretative claims constructed from

5

© AI & Antiquity. Volume 2, Issue 1, 2026.

Actually, generative AI models are not databases. They do not "know" whether a reference exists; they predict text. When prompted to generate or adjust bibliographic entries, they do exactly what they are designed to do: produce something that looks right. The danger lies in our tendency to confuse formal correctness with epistemic reliability, as if the mere presence of a well-formed citation were synonymous with traceability and truth.

In the Humanities, and particularly in Ancient Studies, bibliography is not an accessory. It is the backbone of scholarly argumentation. References are not decorative footnotes; they are claims of verification—signals that knowledge can be traced, contested, and reassessed. When that chain is broken, even unintentionally, credibility collapses, and the reader is deprived of the very mechanism that makes scholarship cumulative rather than rhetorical. What is new in the AI context is the scale and speed at which such errors can propagate. A single fabricated reference, once cited, reformatted, exported to reference managers, or reused in teaching materials, can circulate widely before anyone notices the absence of an original source. This dynamic is amplified by the growing reliance on automated workflows, citation generators, and AI-assisted writing environments that prioritise efficiency over verification. The result is a paradox that we can no longer ignore: tools designed to assist scholarship can, if used uncritically, undermine the very standards they are meant to support.

2. VERIFICATION, DUE PROCESS, AND AI-RELATED ACADEMIC BULLYING

The response to this situation cannot be a return to technological rejection, nor can it rely solely on warnings directed at students. Instead, it requires a collective recalibration of scholarly habits—one that recognises that, in an AI-mediated environment, verification must be treated not as an optional final step but as the core scholarly practice linking responsibility, traceability, and human accountability. In this respect, few approaches are more instructive than the perspective advanced by Solga and Sarwar (2026, this issue) whose emphasis on procedural transparency and evidentiary anchoring foregrounds verification as a constitutive scholarly act rather than a post *hoc* corrective.

First, we must acknowledge that bibliographic vigilance is now a core scholarly skill, on par with source criticism or methodological transparency. Checking references is no longer a cosmetic final operation; it is the epistemic labour that distinguishes human scholarship from automated plausibility. Second, we must resist the temptation to treat AI as a neutral technical assistant. Delegating tasks such as citation completion, bibliography generation, or reference formatting without verification is not a time-saving shortcut; it is a transfer of responsibility to a system that cannot bear it. The historian, archaeologist, or philologist remains accountable for every reference that appears under their name, just as they remain accountable for the evidence, reasoning, and interpretative claims constructed from

those references. Third, this issue must be addressed explicitly in teaching. Students should not only be warned that "AI invents sources", but guided through practical exercises that expose how and why this happens. Asking students to verify AI-generated bibliographies, to track down nonexistent references, or to compare AI outputs with real databases can transform a risk into a powerful learning opportunity, reinforcing critical thinking and digital literacy simultaneously.

For academic journals, the implications are equally serious. Peer review alone is no longer sufficient if reviewers assume that bibliographies are mechanically correct. Editorial workflows must adapt, incorporating explicit checks and fostering a shared awareness that formal plausibility is no guarantee of authenticity. This does not mean increasing bureaucratic burden or policing authors through suspicion. It means recognising that we are operating in a transformed epistemic environment, where traditional signals of reliability can no longer be taken for granted and must be actively verified. Indeed, innovation in teaching and research must go hand in hand with methodological rigour and ethical responsibility. Embracing AI critically does not mean trusting it blindly; it means understanding its limitations and designing practices that compensate for them.

It is essential to make one point absolutely clear: the presence of incorrect (or even nonexistent) bibliographic references should not automatically be framed as academic misconduct, particularly in this transitional phase of widespread AI adoption. In many cases, such errors do not stem from an intention to deceive, fabricate data, or manipulate scholarly discourse, but from the uncritical or partial use of tools whose overall utility far exceeds these specific risks. To reduce AI to a mere channel of misconduct would be both inaccurate and counterproductive, especially when these systems are already embedded in everyday academic practice, from language revision and structural clarity to accessibility and pedagogical design. Moreover, editorial policies that reflexively equate bibliographic hallucinations with fraud risk producing unjust outcomes, especially for early-career researchers, non-native speakers, and scholars working under intense publication pressure. Punitive approaches may, paradoxically, discourage transparency, pushing authors to conceal AI use rather than engage with it openly and responsibly.

A growing number of authors are already encountering the consequences of this transitional moment firsthand. In several reported cases, irregularities in citations, source selection, or textual reconstruction—sometimes linked to AI-assisted workflows, sometimes to misguided methodological shortcuts—have been interpreted not as errors requiring correction but as prima facie evidence of deliberate scientific misconduct. What is striking is not the identification of problems themselves, which is both legitimate and necessary, but the evaluative leap that follows: suspicion is rapidly converted into moral judgement. Irregular or inappropriate citations are treated as proof of intent; methodological weaknesses

are reframed as fabrication; and correctable failures of verification are recoded as indicators of dishonesty.

This progression collapses the distinctions on which responsible scholarly assessment depends. Errors, negligence, and fraud—categories that require different evidentiary thresholds and call for different responses—are flattened into a single presumption of bad faith. The result is a shift from procedural evaluation to moralisation, in which the possibility of clarification, revision, or methodological correction is displaced by the language of sanction. In some instances, authors are not only rejected but threatened with institutional reporting without any prior stage of dialogue, documentation of evidence, or opportunity for response, thereby short-circuiting the processes that ensure proportional and evidence-based judgement.

Documented episodes within recent scholarly communication further illustrate this dynamic. Public attributions of AI authorship, initially advanced on the basis of alleged bibliographic anomalies, have in some cases been shown—following systematic verification of sources, citations, and editorial records—to be factually incorrect and unsupported by evidence. Nevertheless, the original claims circulated rapidly through private communication networks and informal scholarly channels, acquiring credibility through repetition rather than documentation and producing reputational effects without any procedural mechanism for clarification or right of reply. Within competitive and precarious academic environments, such dynamics intersect with factional alignments and reputational economies, where AI suspicion may be mobilised strategically as a tool of exclusion. When unsubstantiated allegations are disseminated in ways that are difficult to contest, detached from formal review, and capable of producing demonstrable professional harm, the boundary between critical evaluation and reputational aggression becomes blurred. In their most problematic form, these practices constitute a mode of academic bullying: not through overt hostility, but through the circulation of unverifiable claims that stigmatise authors while evading the evidentiary standards that govern scholarly critique.

The ethical issue at stake is therefore not only methodological but procedural. Scholarly disagreement requires argument, documentation, and the possibility of response; reputational claims about authorship or integrity require an equivalent evidentiary burden. To advance allegations of AI use without verifiable criteria, and to disseminate them through informal channels that preclude correction, risks approximating the dynamics of defamation in functional terms, even when framed as methodological concern. Detached from transparent processes, provisional suspicions can harden into durable reputational narratives that are resistant to later clarification. These cases underscore the necessity of shared protocols that distinguish clearly between documented irregularity, correctable error, and demonstrable misconduct, and that subject both texts and accusations to equivalent evidentiary standards as a condition of scholarly integrity.

Instead, academic publishers and editorial boards must recognise bibliographic fabrication by AI as a systemic by-product of a rapidly evolving technology, not as evidence of individual ethical failure *per se*. The appropriate response is not sanction, but shared responsibility: clearer guidelines on acceptable AI use, explicit verification protocols, and constructive communication between editors, reviewers, and authors when such issues arise. This position does not imply lowering standards. On the contrary, it reinforces them. Scholarly rigour is best protected not through suspicion and punishment, but through education, awareness, and procedural adaptation—much as the academic community once learned to integrate digital databases, reference managers, and online archives without criminalising their early misuses. Indeed, recent reporting has made clear that this is no longer an anecdotal classroom issue. Cases are emerging in which real academic outlets cite references that appear to have been fabricated or laundered through AI-mediated workflows, blurring the line between error, automation, and institutional negligence. Bibliographic hallucinations are thus migrating from student submissions into the wider circulation of scholarly writing—precisely because they imitate our conventions so well.

Reactions across scholarly networks are instructive. Some researchers return to traditional, verifiable search practices; others rely on specialised discovery platforms or AI tools designed for bibliographic validation; still others read the phenomenon as a symptom of structural pressures such as citation inflation, paywalls, and editorial workflows that do not systematically verify references. What matters editorially is not which response is "correct", but what they collectively signal: scholarship must now actively defend itself against the production of plausible but unverified knowledge. This requires not only attention to student practice but the refinement of institutional norms, editorial procedures, and everyday research habits.

The task is not to reject AI, but to embed verification as a non-negotiable scholarly standard, ensuring that the ease of generating academically styled text does not weaken the chain of traceability on which humanistic knowledge depends. It is against this backdrop of uncertainty, eroded trust, and renewed demands for evidence that the present issue must be read. Volume 2, Issue 1 approaches artificial intelligence neither as a neutral innovation nor as a threat to be contained, but as a field of practice in which mediation, method, responsibility, and inclusion must be actively negotiated. The contributions gathered here offer not technical fixes but concrete pedagogical, methodological, and historiographical strategies for working with AI without relinquishing scholarly rigour.

3. HIGHLIGHTS OF THIS ISSUE

Taking into account all of the above, *Volume 2, Issue 1* of *AI & Antiquity* has been conceived not as a loose collection of contributions, but as a deliberately structured

itinerary, one that moves from the classroom to research practice, from methodological experimentation to cultural responsibility, and ultimately to the ethical question that remains beneath every technological debate: what is the purpose of knowledge, and whose voices does it serve?

The issue opens with a first thematic block centred on AI as a mediator of knowledge, with a clear emphasis on the classroom as the primary laboratory of transformation. In "From Classical Sources to Artificial Intelligence: *Notebook LM* as a Cognitive Mediator in University Teaching of Ancient History", Francisco Javier Catalán González offers an experimental, source-based case study built around a controlled *corpus* of Greco-Roman narratives on the foundation of Rome (Livy, Dionysius of Halicarnassus, Cassius Dio, and Velleius Paterculus, using in this case the Gredos editions). Working within a closed-document environment, the article tests how *Notebook LM* can process and synthesise primary materials and then transform them into university-oriented learning artefacts—summaries, concept maps, flashcards, and multimodal outputs (audio and video)—designed to support cognitive structuring, retention, and active study. Crucially, the study does not present AI as an interpretative substitute, but as a scaffold for engaging with dense textual corpora; at the same time, it identifies two pedagogically significant risks—discursive homogenisation and limited philological sensitivity—thereby reinforcing the issue's wider argument that AI can improve accessibility and comprehension only under explicit human supervision and source-aware teaching design.

This pedagogical entry point is then expanded and reframed through "Let's Chat About Archaeology: Responsible and Thoughtful Use of AI Tools in the Classroom, A Case Study" by Yusi Liu, Daniel D'Elia, and Rocco Palermo. Based on their Bryn Mawr course, the authors design an assignment in which students work both as AI-assisted writers and as editors who annotate, fact-check, and revise chatbot outputs using citations and tracked changes. By linking prompt design, collaborative verification, and the identification of inaccuracies and bias to the wider problem of pseudo-archaeological misinformation, the study provides a transferable model for integrating AI into teaching while foregrounding accountability, transparency, and critical evaluation.

From the classroom, the issue then moves into research practice with "Evaluating Generative AI in Historical Research: A Comparative Study on Identifying Primary Source Evidence in Ancient History" by Raymond S. Solga and Mohammed J. Sarwar. Through a paired case-study design, the authors contrast four human-led investigations with four AI-assisted inquiries conducted across GPT-4, Claude 2, Gemini, and Perplexity. The comparison is structured around explicitly historiographical criteria—temporal framing, provenance transparency, genre differentiation, linguistic fidelity, and evidentiary reliability—making visible both the exploratory speed of generative systems and their persistent weaknesses in verification, manuscript lineage, and genre control.

The next step in the issue's arc functions as a conceptual hinge. "Algorithmic Memory: Towards Reflexive Authenticity in Cultural Heritage" by Menna Salah shifts the focus from research method to the public life of the past. The article argues that artificial intelligence is not merely a tool of preservation but a force that reshapes how cultural memory is curated, legitimised, and shared. Through the notion of *reflexive authenticity*, authenticity is redefined as transparency rather than curatorial authority as participatory rather than hierarchical.

"Recovering the Voices of Silence in Ancient Historiography: A Re-reading of the *Shiji* through the Lens of Inclusive Artificial Intelligence" by Samandar Ruziboev and Noyibjon Khudoyorov brings the volume to its most explicitly historiographical horizon. Focusing on Sima Qian's *Shiji* (ca. 145–86 BCE) within the political context of the early Han empire, the article re-reads silence and diplomacy not as narrative absences but as structuring features of historical writing. AI is employed not as an interpretive authority but as an analytical instrument capable of tracing patterns that elude close reading, opening a methodological space in which questions of power, language, and representation can be reconsidered. The contribution thus moves the debate from procedures to meaning, asking how inclusive and reflexive approaches reshape whose voices are transformed in ancient historiography.

After the peer-reviewed articles, the volume turns to a section that brings these debates into the sphere of Public History and contemporary narratives of Antiquity. The contributions by Iban Martín, Patricia González, and Mario Agudo, presented in the seminar on *Ancient World Today* organised with Marc Mendoza at the UAB (27 November 2025), show how the questions raised throughout the issue move beyond the classroom and research settings into public discourse, where Antiquity is told in ways that shape contemporary understanding. What emerges is not only a methodological concern, but the persistent unevenness of the archive itself—those presences that survive only in fragments and those that never entered the record, including many women whose experiences remain largely inaccessible. Making those absences visible requires not only new tools, but renewed commitments to verification, proportionality, and the rejection of bullying practices in scholarly evaluation.

Bellaterra (Barcelona), February 2026

BIBLIOGRAPHY

Liu, Y., D'Elia, D. and Palermo, R. (2026) 'Let's Chat About Archaeology: Responsible and Thoughtful Use of AI Tools in the Classroom, A Case Study', *AI & Antiquity*, 2(1), this issue.

Solga, R. and Sarwar, M. J. (2026) 'Evaluating Generative AI in Historical Research: A Comparative Study on Identifying Primary Source Evidence in Ancient History', *AI & Antiquity*, 2(1), this issue.

Walters, W. H. and Wilder, E. I. (2023) 'Fabrication and errors in the bibliographic citations generated by ChatGPT', *Scientific Reports*, 13, article 14045. doi: 10.1038/s41598-023-41032-5.

De las fuentes clásicas a la Inteligencia Artificial: *Notebook LM* como mediador cognitivo en la enseñanza universitaria de Historia Antigua

From Classical Sources to Artificial Intelligence: Notebook LM *as a Cognitive Mediator in University Teaching of Ancient History*

Francisco Javier Catalán González

Instituto Arqueológico de Mérida (IAM)
CSIC-Junta de Extremadura
javier.catalan@iam.csic.es
0000-0003-2810-3477

RESUMEN — El presente estudio analiza la utilidad de la herramienta *Notebook LM*, desarrollada por Google, como recurso didáctico en la enseñanza universitaria de Historia Antigua, tomando como caso de estudio la fundación de Roma. A partir de un corpus de textos grecolatinos de la editorial Gredos (Tito Livio, Dionisio de Halicarnaso, Dion Casio y Veleyo Patérculo), la investigación evalúa la capacidad de la inteligencia artificial (IA) para procesar, sintetizar y generar materiales educativos adaptados al contexto académico. El trabajo adopta una metodología experimental basada en la carga de documentos en un entorno cerrado, la configuración de un *prompt* orientado a la producción de contenidos universitarios y la posterior evaluación cualitativa de los resultados. Los materiales generados, resúmenes, mapas conceptuales, tarjetas didácticas, audios y vídeos evidencian que *Notebook LM* favorece la estructuración cognitiva y la retención del conocimiento, promoviendo un aprendizaje multimodal y activo. No obstante, se advierte una tendencia a la homogeneización del discurso histórico y una limitada sensibilidad filológica, lo que subraya la necesidad de una supervisión humana. En conjunto, la IA se revela como un instrumento complementario y mediador cognitivo, capaz de mejorar la accesibilidad y la comprensión del conocimiento histórico, sin sustituir la interpretación crítica propia de las Humanidades.

PALABRAS CLAVE — Inteligencia Artificial, Didáctica en Historia Antigua, Teoría de la Educación, mediación cognitiva.

ABSTRACT — This study examines the usefulness of *Notebook LM*, developed by Google, as an educational tool in university-level Ancient History, focusing on the case study of the foundation of Rome. Based on a corpus of Greco-Roman texts from

the Gredos collection (Titus Livy, Dionysius of Halicarnassus, Cassius Dio, and Velleius Paterculus), the research evaluates the capacity of artificial intelligence (AI) to process, synthesize, and generate academic learning materials. The methodology followed an experimental design involving the upload of primary sources into a closed environment, the configuration of a university-level *prompt*, and the qualitative evaluation of the produced outputs. The materials generated, summaries, concept maps, flashcards, audio and video resources demonstrate that *Notebook LM* enhances cognitive structuring and knowledge retention, fostering multimodal and active learning. However, the tool also tends to homogenize historical discourse and shows limited philological sensitivity, highlighting the need for human supervision. Overall, AI emerges as a complementary and cognitive mediation tool, capable of improving accessibility and comprehension in historical learning without replacing the critical interpretation central to the Humanities.

1. INTRODUCCIÓN

1.1 CONTEXTUALIZACIÓN DEL USO DE LA INTELIGENCIA ARTIFICIAL EN EL APRENDIZAJE DE LA HISTORIA

En los últimos años, el desarrollo de la inteligencia artificial (IA) ha transformado de manera significativa los procesos de enseñanza-aprendizaje en el ámbito universitario. La irrupción de sistemas de procesamiento del lenguaje natural, aprendizaje automático y generación de contenidos digitales ha abierto nuevas posibilidades para la personalización educativa y la creación de materiales didácticos dinámicos (Luckin, 2022; Zawacki-Richter *et al.*, 2019). En este contexto, las disciplinas de humanidades, y particularmente la Historia, han comenzado a explorar el potencial de estas herramientas como medio para fomentar la comprensión crítica de los contenidos, el análisis de fuentes y la autonomía del estudiante (Cope and Kalantzis, 2023).

La enseñanza de la Historia Antigua presenta retos específicos relacionados con la distancia temporal, la interpretación de fuentes primarias y la necesidad de contextualizar los procesos históricos en estructuras sociales y culturales complejas. A ello se suma la dificultad de mantener la motivación del alumnado ante un *corpus* textual denso o de difícil comprensión para estudiantes de los primeros años del Grado. Frente a estos desafíos, las tecnologías basadas en IA ofrecen oportunidades inéditas para reconfigurar el acceso a la información histórica mediante sistemas capaces de resumir, relacionar y reformular contenidos con precisión y coherencia didáctica (Holmes *et al.*, 2021; García-Peñalvo, 2024).

El avance reciente de modelos de lenguaje de gran escala (*LLMs*, por sus siglas en inglés) ha permitido la integración de asistentes educativos que pueden procesar

grandes volúmenes de información y generar materiales adaptados a distintos niveles de comprensión. Entre estas herramientas emergentes destaca *Notebook LM*, desarrollada por Google, concebida como un entorno de aprendizaje asistido por IA que permite importar un conjunto de documentos y generar, a partir de ellos, material educativo diverso: resúmenes, mapas conceptuales, cuestionarios, y explicaciones en formato textual, auditivo o audiovisual (Google, 2024). Este enfoque posibilita una experiencia de estudio interactiva, centrada en la comprensión y retención del conocimiento, en lugar de en la mera memorización de datos.

1.2 Objetivos del estudio y justificación

El presente trabajo tiene como objetivo analizar la utilidad de *Notebook LM* en el aprendizaje de contenidos de la asignatura de Historia Antigua II, tomando como caso de estudio la temática de la fundación de Roma. Se ha buscado evaluar la capacidad de esta herramienta para generar materiales de apoyo al estudio a partir de fuentes históricas primarias, y determinar su pertinencia pedagógica en el contexto universitario. El estudio pretende, asimismo, explorar la interacción entre IA y conocimiento humanístico, examinando en qué medida la automatización de la síntesis y estructuración del contenido contribuye al desarrollo de habilidades analíticas en los estudiantes.

La elección del tema, la fundación de Roma responde tanto a su relevancia en el temario de Historia Antigua II como a su potencial como campo de experimentación, dado que cuenta con un amplio número de fuentes grecolatinas que abordan el mito fundacional desde perspectivas diversas: desde los relatos de Tito Livio, Dionisio de Halicarnaso, entre muchos otros. El uso de este *corpus* textual ofrece un contexto idóneo para analizar la capacidad de *Notebook LM* de integrar, sintetizar y representar información histórica compleja a partir de fuentes heterogéneas.

2. Marco teórico

2.1 La Historia Antigua como disciplina y sus retos pedagógicos

La Historia Antigua constituye una de las áreas fundamentales dentro de la carrera de Historia. Su enseñanza universitaria implica un enfoque crítico del pasado que combina el análisis de textos literarios, epigráficos y arqueológicos con la interpretación de procesos políticos, sociales y culturales (Momigliano, 1990; Finley, 1985). En este sentido, la formación en Historia Antigua no se reduce a la adquisición de datos cronológicos, sino que requiere desarrollar competencias orientadas a la comprensión del tema en base a los textos históricos que se suelen plantear en clase.

Sin embargo, el contexto pedagógico actual plantea importantes desafíos. La creciente digitalización del aprendizaje ha modificado la manera en que los estudiantes se aproximan a las fuentes históricas. Frente a los modelos tradicionales

de enseñanza basados en la lectura lineal y el comentario guiado, las nuevas generaciones interactúan con los contenidos mediante entornos digitales que priorizan la inmediatez, la visualidad y la fragmentación informativa (Carr, 2020; Berti, 2023). Este cambio en los hábitos de estudio ha llevado a reflexionar sobre la necesidad de renovar los métodos docentes para fomentar una comprensión activa y sostenida del conocimiento histórico (Prensky, 2011).

En este contexto, la aplicación de tecnologías basadas en inteligencia artificial se perfila como una herramienta de apoyo al aprendizaje histórico, no para sustituir la interpretación crítica, sino para facilitar el acceso a los textos y promover una lectura guiada e interactiva. Las plataformas que permiten el tratamiento automatizado de fuentes documentales ofrecen nuevas posibilidades para generar síntesis estructuradas, mapas conceptuales o esquemas cronológicos, ayudando a los estudiantes a organizar y retener la información sin perder el vínculo con la reflexión historiográfica (Cano, 2023).

2.2 Inteligencia Artificial y aprendizaje universitario

La incorporación de la inteligencia artificial en el ámbito educativo ha sido objeto de creciente interés académico durante la última década. La IA aplicada a la enseñanza se define como el conjunto de sistemas capaces de analizar, generar o adaptar contenidos con el fin de mejorar los procesos de aprendizaje y evaluación (Luckin, 2022; Holmes *et al.*, 2021). En el contexto universitario, su uso se ha orientado tanto a la personalización del aprendizaje como al apoyo a la investigación, mediante asistentes conversacionales, tutores virtuales y entornos de análisis semántico.

El potencial de estas tecnologías radica en su capacidad para transformar datos en conocimiento estructurado, ofreciendo explicaciones, resúmenes y conexiones entre conceptos de manera contextualizada. De acuerdo con Zawacki-Richter *et al.* (2019), las aplicaciones de IA en educación superior se pueden agrupar en tres dimensiones principales: apoyo al estudiante, asistencia al profesorado en la evaluación y diseño curricular, y análisis predictivo del rendimiento académico (*Figura 1*). Sin embargo, en el ámbito de las Humanidades, su incorporación ha sido más lenta, debido tanto a la naturaleza interpretativa del conocimiento histórico como a los debates epistemológicos que suscita la mediación tecnológica en la construcción del saber (Cope and Kalantzis, 2023).

Desde una perspectiva didáctica, diversos autores han señalado que la IA no debe concebirse como un sustituto de la enseñanza tradicional, sino como un medio complementario que potencie la autonomía cognitiva del estudiante (García-Peñalvo, 2024). La interacción con sistemas inteligentes permite simular escenarios de aprendizaje guiado, generar retroalimentación inmediata y adaptar los contenidos a distintos niveles de comprensión. De esta forma, la IA puede contribuir a superar una de las limitaciones habituales en la enseñanza universitaria de

Historia: la dificultad de articular la lectura de textos clásicos con la comprensión del contexto histórico y la reflexión crítica.

2.3 *Notebook LM* COMO ENTORNO DE APRENDIZAJE MEDIADO POR IA

Notebook LM representa una de las propuestas más recientes dentro del ámbito de la inteligencia artificial educativa. Desarrollada por Google, esta herramienta utiliza modelos de lenguaje avanzados (*LLMs*) para analizar, sintetizar y relacionar información procedente de un conjunto de documentos proporcionados por el usuario (Google, 2024). Su principal característica radica en la posibilidad de construir un "cuaderno digital" a partir de una biblioteca de fuentes, lo que permite generar materiales de aprendizaje adaptados al contenido y al nivel de comprensión deseado.

La interfaz de *Notebook LM* (*Figura 2*) integra funciones que permiten resumir textos extensos, elaborar explicaciones temáticas, formular preguntas tipo test, y producir versiones en formato audio o video de los materiales elaborados. A diferencia de otros entornos de IA conversacional, como *ChatGPT* o *Claude*, *Notebook LM* opera sobre una base documental cerrada, definida por el propio usuario, lo que garantiza una mayor trazabilidad y control de las fuentes utilizadas (*Google Research*, 2024). Este enfoque resulta especialmente pertinente para la enseñanza de Historia, donde la fiabilidad y procedencia del material constituyen aspectos esenciales de la práctica académica.

Desde una perspectiva pedagógica, *Notebook LM* puede considerarse una herramienta de mediación cognitiva que articula tres funciones principales: Organización del conocimiento, mediante la síntesis y jerarquización de la información contenida en múltiples documentos. Representación multimodal, al transformar los contenidos en materiales accesibles en distintos formatos (textual, auditivo y visual). Y el aprendizaje interactivo, al permitir que el estudiante explore, formule preguntas y obtenga explicaciones personalizadas a partir del documento seleccionado.

Diversos estudios recientes han comenzado a señalar el valor de los entornos de IA centrados en documentos como instrumentos de aprendizaje activo y reflexión guiada (Baker and Smith, 2023). En el caso del estudio aquí presentado, el programa fue utilizado para procesar un *corpus* de obras de historiadores grecolatinos, con el fin de generar un conjunto de materiales didácticos sobre la fundación de Roma. Este enfoque experimental busca evaluar la capacidad de la IA para articular conocimiento histórico a partir de fuentes primarias, manteniendo la precisión conceptual y la coherencia narrativa.

3. METODOLOGÍA

3.1 SELECCIÓN DE FUENTES PRIMARIAS

La fase inicial del estudio consistió en la compilación de un conjunto de obras de la Editorial Gredos, una de las más reconocidas en el ámbito de las traducciones críticas de autores grecolatinos al castellano, que previamente habían sido digitalizados y convertidos a formato PDF. Se seleccionaron 17 documentos (la versión gratuita permite hasta 50).

Entre los autores incluidos se encuentran Tito Livio, con 8 documentos de *Historia de Roma desde su fundación*; Dionisio de Halicarnaso, con *Historia Antigua de Roma*, con un total de 4 documentos; Dion Casio, *Historia de Roma*, 4 volúmenes; Veleyo Patérculo, *Historia Romana*. Cabe destacar que, en el caso de Tito Livio, los "ocho documentos" a los que se hace referencia corresponden al formato de la edición de Gredos de *Ab Urbe Condita*, y no al total de libros que integraban originalmente la obra. El criterio de selección se basó en tres principios:

- Relevancia temática: inclusión de textos directamente relacionados con los relatos fundacionales, las genealogías míticas de Eneas y Rómulo, o las primeras instituciones del Estado romano.
- Representatividad historiográfica: equilibrio entre autores que aportaran perspectivas complementarias sobre los mismos acontecimientos.
- Accesibilidad textual: uso de traducciones modernas y anotadas que garantizaran la legibilidad y coherencia terminológica dentro del *corpus*.

Cada documento fue digitalizado o convertido a formato PDF y posteriormente cargado en la plataforma *Notebook LM* como base de conocimiento. Este conjunto de textos constituyó el material de referencia sobre el cual el modelo de lenguaje generó los distintos productos educativos.

3.2 DISEÑO DEL EXPERIMENTO DIDÁCTICO

El diseño metodológico adoptó un enfoque experimental descriptivo, orientado a explorar las potencialidades de *Notebook LM* como herramienta de generación de materiales didácticos. El proceso se articuló en tres fases principales:

- Carga y organización: los documentos fueron introducidos en *Notebook LM* como fuentes base, creando un entorno documental cerrado y homogéneo. La herramienta permitió vincular internamente los textos, posibilitando que el modelo estableciera relaciones entre autores y pasajes.
- Definición del *Prompt*: se formuló una instrucción dirigida al modelo con el siguiente enunciado:

  "Utiliza un lenguaje formal, adaptado al ámbito de comprensión universitaria. Elabora materiales de apoyo para el estudio de la fundación de

Roma. Remarca el uso de elementos mitológicos y elementos históricos. Busca en las notas de los traductores datos relativos a la utilización del mito como refuerzo político. Evita reiteraciones".

Una vez configurado el entorno y establecidas las directrices de estilo, se procedió a la generación de distintos tipos de materiales, con el objetivo de evaluar la versatilidad didáctica de la *IA*. Entre los productos elaborados destacan:

- Resumen de audio: *Notebook LM* permitió convertir los contenidos en un formato de audio narrativo, facilitando la asimilación pasiva de información. Este recurso resulta especialmente útil en el aprendizaje ubicuo, permitiendo al estudiante reforzar conceptos durante desplazamientos o actividades no sedentarias.
- Resumen de video: La interfaz generó una versión audiovisual condensada, que sintetiza los aspectos más relevantes de la fundación de Roma mediante una narración secuencial apoyada en texto y voz. Su formato es compatible con plataformas educativas y entornos virtuales de aprendizaje (EVA) (*Figura 3*).
- Mapa mental: A partir de la interpretación de los textos, *Notebook LM* estructuró un mapa conceptual que muestra las relaciones entre mitos fundacionales, personajes y elementos. Este tipo de representación visual es especialmente eficaz de cara a asimilar el temario.
- Informes temáticos: Se generaron informes analíticos en formato textual, donde la IA sintetizó los argumentos de las distintas fuentes, comparando perspectivas historiográficas y señalando discrepancias o convergencias interpretativas.
- Tarjetas didácticas: Se crearon *flashcards* con definiciones breves, citas destacadas y preguntas de repaso, permitiendo la aplicación de técnicas de *spaced repetition* para el refuerzo de la memoria a largo plazo (*Figura 4*).

El procedimiento se llevó a cabo de manera iterativa, solicitando a la IA que reelaborara o ampliara ciertos apartados cuando se detectaban ambigüedades, errores de interpretación o exceso de simplificación. En todo momento se procuró mantener la fidelidad al contenido original de las fuentes, sin introducir información ajena a los documentos previamente cargados.

3.3 CRITERIOS DE EVALUACIÓN DE LOS MATERIALES PRODUCIDOS

La evaluación de los materiales generados por *Notebook LM* se llevó a cabo mediante un análisis cualitativo estructurado, centrado en cuatro dimensiones fundamentales:

- Exactitud histórica: se verificó la fidelidad del contenido respecto a las fuentes primarias. Se analizaron los casos en que la IA sintetizaba o

reformulaba información, comprobando si mantenía la coherencia con el texto original.

- Rigor académico y terminológico: se revisó el uso de terminología histórica, la precisión conceptual y la adecuación del lenguaje al nivel universitario solicitado en el *Prompt.*
- Pertinencia didáctica: se valoró la utilidad de los materiales para el aprendizaje, su claridad expositiva, y su capacidad para fomentar la comprensión de procesos complejos.
- Coherencia multimodal: se compararon los resultados entre los distintos formatos (texto, audio, vídeo, mapas conceptuales, cuestionarios) para determinar si existía correspondencia narrativa y conceptual entre ellos.

El análisis se complementó con la observación de posibles errores sistemáticos, como la tendencia a la generalización excesiva, la omisión de nombres propios o la simplificación de estructuras narrativas. Los materiales se clasificaron según su nivel de precisión y relevancia pedagógica, empleando una escala de tres niveles: *adecuado*, *parcialmente adecuado* y *no adecuado*.

3.4 LIMITACIONES DEL ESTUDIO

Es importante señalar algunas limitaciones metodológicas inherentes al uso de herramientas de inteligencia artificial en el ámbito de la investigación histórica. En primer lugar, la dependencia de la documentación: al estar el sistema limitado a los documentos cargados, su comprensión se circunscribe a ese contexto textual y no puede contrastar con fuentes no incluidas. En segundo lugar, la falta de sensibilidad filológica: aunque el modelo es capaz de sintetizar y reorganizar información, tiende a uniformar matices lingüísticos y conceptuales propios de los textos clásicos (González, 2024). Finalmente, la dificultad de evaluación cuantitativa: la valoración de la calidad interpretativa y pedagógica de los resultados mantiene un componente subjetivo inevitable.

4. EVALUACIÓN CUALITATIVA DEL PROCESO

Desde el punto de vista pedagógico, la experiencia evidenció la capacidad de *Notebook LM* para actuar como un asistente de aprendizaje contextualizado, capaz de integrar fuentes primarias extensas y generar productos derivados con alto grado de coherencia interna. El sistema demostró ser especialmente eficaz en la identificación de conceptos clave, en la estructuración jerárquica de la información y en la elaboración de materiales de estudio diversificados.

No obstante, se detectaron algunas limitaciones inherentes al procesamiento automático del lenguaje histórico, cuestiones interpretativas que se analizan con mayor detalle en el apartado de discusión.

5. RESULTADOS PRELIMINARES

El análisis de los materiales generados permite afirmar que el uso de *Notebook LM* en la enseñanza de la Historia Antigua favorece una mayor accesibilidad cognitiva y una mejora en la organización del conocimiento. Los estudiantes pueden interactuar con la IA para aclarar dudas, solicitar resúmenes parciales o comparar perspectivas, generando una dinámica de estudio más activa y personalizada.

Asimismo, la posibilidad de integrar diversos formatos (audio, texto, mapa mental, cuestionarios) contribuye a la multimodalidad del aprendizaje, aspecto fundamental en las pedagogías contemporáneas. Este enfoque promueve la retención mediante la exposición a diferentes estímulos sensoriales y cognitivos, ajustándose a las necesidades individuales del estudiante.

6. DISCUSIÓN

6.1 LA INTELIGENCIA ARTIFICIAL COMO HERRAMIENTA EN LA ENSEÑANZA DE LA HISTORIA ANTIGUA

Los resultados obtenidos confirman que *Notebook LM* puede desempeñar un papel significativo en la enseñanza en Historia Antigua. Su capacidad para procesar grandes volúmenes de información textual, identificar conceptos centrales y establecer relaciones jerárquicas entre ellos convierte a la inteligencia artificial en un elemento de apoyo eficaz dentro del proceso de aprendizaje.

La experiencia realizada mostró que el modelo es capaz de generar materiales coherentes, tanto en el plano conceptual como en el discursivo. El sistema organizó la información en torno a los principales núcleos narrativos, manteniendo un tono formal y académico, acorde con las instrucciones del *Prompt*. Esta respuesta estructurada indica que, cuando se trabaja con un conjunto de textos delimitado y homogéneo, la IA puede comportarse como un intérprete de segundo orden. En este sentido, *Notebook LM* demuestra una ventaja respecto a otros modelos de lenguaje de propósito general, tal propiedad refuerza la confianza en el uso académico del sistema, siempre que se mantenga la supervisión crítica del docente o investigador (Fernández and Molina, 2024).

Desde el punto de vista pedagógico, los resultados podrían suponer una mejora en la organización del conocimiento histórico, favoreciendo la asimilación de conceptos, tal como sugiere la teoría del aprendizaje significativo de Ausubel (1968). El mapa mental elaborado por la IA permitió representar de forma visual las conexiones entre mito, territorio, personajes y legitimación del poder, lo cual facilitó la comprensión de un proceso histórico complejo mediante esquemas relacionales. Las tarjetas didácticas, por su parte, pueden contribuir al refuerzo memorístico mediante repetición espaciada, mientras que los resúmenes de audio y vídeo introdujeron una dimensión multimodal del aprendizaje, coherente con los enfoques actuales de educación digital (Mayer, 2021). Asimismo, el cuestionario

19

final mostró una adecuada correspondencia entre los contenidos esenciales y los niveles de dificultad, permitiendo su uso como instrumento de autoevaluación formativa.

6.2 Limitaciones observadas y riesgos epistemológicos

No obstante, la experimentación reveló una serie de limitaciones inherentes al uso de inteligencia artificial en contextos humanísticos. En primer lugar, se observó una tendencia a la homogeneización del discurso, donde las divergencias entre autores clásicos tienden a diluirse en favor de una narrativa coherente pero simplificada. Por ejemplo, en los apartados de resumen de audio y vídeo, donde las diferencias entre la visión moralizante de Tito Livio y el enfoque analítico de Dionisio de Halicarnaso se reducen a una exposición neutral que pierde parte de su riqueza interpretativa.

En segundo lugar, la inteligencia artificial mostró dificultades para abordar conceptos polisémicos y culturalmente contextualizados, como *virtus*, *pietas* o *auctoritas*. Si bien la traducción formal de estos términos resulta correcta, la herramienta tiende a uniformarlos desde el punto de vista semántico, desatendiendo tanto su evolución histórica como su función ideológica en los distintos autores y contextos de producción. Esta limitación se manifestó de manera especialmente clara en el podcast y en los audios del video-resumen, donde la ausencia de matizaciones interpretativas resulta evidente desde una perspectiva crítica. En consecuencia, se refuerza la necesidad de una mediación docente que complemente y supervise las interpretaciones generadas por la IA, garantizando una lectura histórica atenta a los matices ideológicos y culturales de las fuentes y evitando su uso como sustituto del análisis historiográfico.

7. Consideraciones sobre la integración institucional de la IA en Humanidades

Los resultados de esta experiencia apuntan hacia una posible integración curricular futura de herramientas de inteligencia artificial en los programas de enseñanza universitaria de Humanidades. No obstante, conviene subrayar que el presente trabajo se sitúa en un plano fundamentalmente teórico y metodológico, ya que no ha sido posible llevar a cabo una aplicación directa de los materiales generados en el aula ni una recogida sistemática de la valoración del estudiantado y del profesorado. La implementación práctica de estos recursos, así como su evaluación empírica en contextos docentes reales, se plantea por tanto como una línea prioritaria de investigación futura.

Desde esta perspectiva, el enfoque adoptado coincide con las directrices del *European Digital Education Framework* (Comisión Europea, 2024), que recomienda fomentar la competencia digital docente y el uso ético de la inteligencia artificial como apoyo al aprendizaje personalizado. En el ámbito específico de la Historia Antigua, la IA ofrece un potencial significativo para la mediación didáctica del

patrimonio textual clásico, facilitando su acceso mediante interfaces interactivas y formatos multimodales adaptados a los hábitos de aprendizaje contemporáneos. La consolidación de estas prácticas requerirá, sin embargo, el establecimiento de protocolos institucionales claros en relación con la autoría, la fiabilidad y la evaluación del material generado por IA, así como una reflexión epistemológica sobre la noción misma de "interpretación" en entornos digitales.

En conjunto, *Notebook* LM puede considerarse una herramienta pertinente para la enseñanza y el estudio de la Historia Antigua, siempre que su uso se enmarque en una supervisión académica rigurosa y responda a objetivos didácticos claramente definidos. Su principal aportación reside en la capacidad para transformar grandes volúmenes de información textual en conocimiento estructurado y accesible, mientras que su principal riesgo radica en la posible pérdida de profundidad interpretativa. El equilibrio entre automatización y reflexión crítica emerge, por tanto, como el eje central de su aplicación pedagógica.

FUENTES

Dion Casio (2004). *Historia romana. Libros XXXVI-XLV*, Trad. y notas de J.M. Candau Morón y M.L. Puertas Castaños, *Biblioteca Clásica Gredos* 326, Gredos, Madrid.

Dionisio de Halicarnaso. (1984). *Historia antigua de Roma. Libros I-III* (E. Jiménez and E. Sánchez, Trads.; introducción de D. Plácido). Editorial Gredos. (Biblioteca Clásica Gredos, 73). ISBN 84-249-0950-X.

Tito Livio (1990). *Historia de Roma desde su fundación*, introducción por A. Sierra, traducción por J.A. Villar Vidal, Biblioteca Clásica Gredos 144, Madrid.

Veleyo Patérculo. (2001). *Historia romana* (E. del Barrio Sanz, Trad.). Editorial Gredos. (Biblioteca Clásica Gredos, 284). ISBN 84-249-2284-0.

BIBLIOGRAFÍA

Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. Holt, Rinehart and Winston.

Baker, T., and Smith, L. (2023). *AI literacy and higher education: Rethinking learning through artificial intelligence*. Routledge.

Berti, A. (2023). *Digital humanities and historical knowledge: Challenges and opportunities*. Springer.

Cano, M. (2023). Nuevas tecnologías y enseñanza de la Historia: desafíos epistemológicos en la era digital. *Revista de Humanidades Digitales, 12*(2), 45–67.

Carr, N. (2020). *The shallows: What the internet is doing to our brains*. W. W. Norton and Company. Consultado el 18/09/2025.

Comisión Europea. (2024). *European Digital Education Framework: Guidelines for Digital Competence in Higher Education*. Bruselas: European Commission. Consultado el 18/09/2025.

Cope, B., and Kalantzis, M. (2023). *Artificial intelligence and the future of teaching and learning*. University of Illinois Press.

Fernández, L., and Molina, R. (2024). Inteligencia artificial y educación universitaria: mediadores tecnológicos en Humanidades. *Revista de Educación y Tecnología, 12*(1), 45–67.

Finley, M. I. (1985). *The use and abuse of history*. Penguin Books.

García-Peñalvo, F. J. (2024). Inteligencia artificial generativa y educación: Un análisis desde múltiples perspectivas. *Education in the Knowledge Society (EKS)*, *25*, e31942. https://doi.org/10.14201/eks.31942

González, M. (2024). IA y fuentes clásicas: desafíos interpretativos en Humanidades Digitales. *Cuadernos de Filología Clásica, 41*, 77–98.

Google. (2024). *Notebook LM: A new way to learn with your notes. Google AI Blog*. Recuperado de https://ai.googleblog.com. Consultado el 22/09/2025.

Google Research. (2024). *Notebook LM: Technical overview*. Recuperado de https://research.google. Consultado el 22/09/2025.

Holmes, W., Bialik, M., and Fadel, C. (2021). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Luckin, R. (2022). *Machine learning and human intelligence: The future of education for the 21st century*. UCL Institute of Education Press.

Mayer, R. E. (2021). *Multimedia learning* (3rd ed.). Cambridge University Press.

Momigliano, A. (1990). *The classical foundations of modern historiography*. University of California Press.

Prensky, M. (2011). Digital natives, digital immigrants. *MCB University Press*.

UNESCO. (2023). *Artificial intelligence in education: Ethical and responsible use*. París: UNESCO. Consultado el 20/10/2025.

Zawacki-Richter, O., Marín, V. I., Bond, M., and Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education, 16*(1), 1–27.

FIGURAS
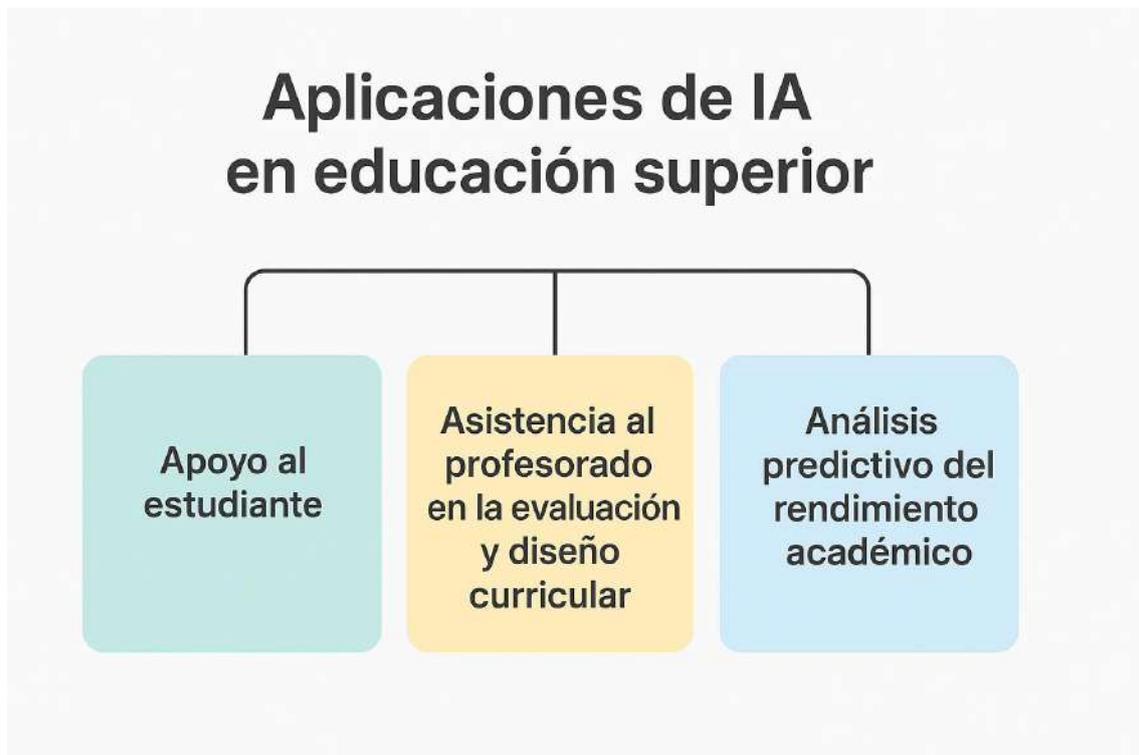
Figura 1. Esquema sintético de las principales aplicaciones de la inteligencia artificial en educación superior, estructuradas en tres dimensiones: apoyo al estudiante, asistencia al profesorado en procesos de evaluación y diseño curricular, y análisis predictivo del rendimiento académico. *Canva*, elaboración propia.
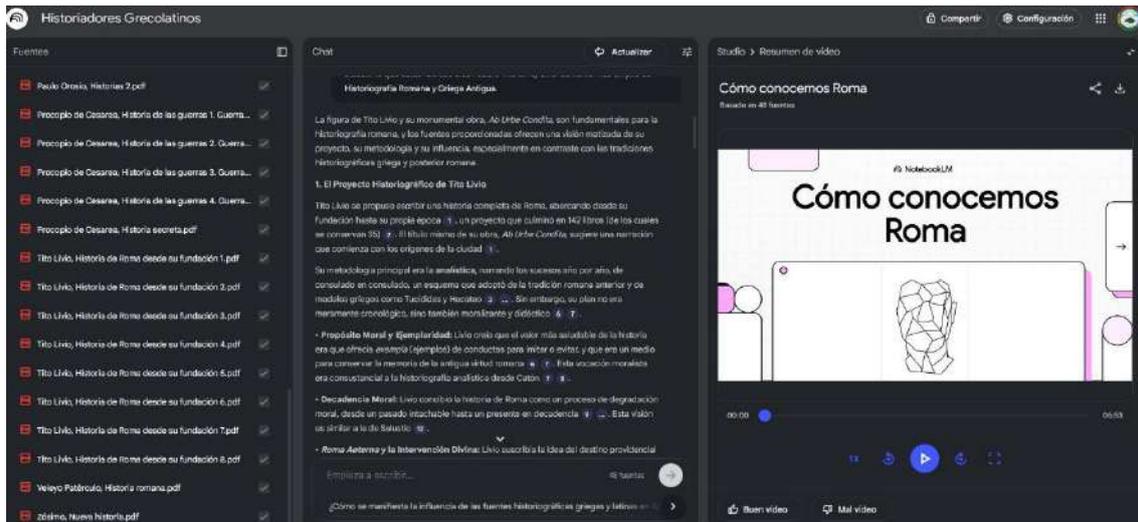
Figura 2. Interfaz principal de *Notebook LM* tras la carga del corpus de fuentes grecolatinas. En la imagen se aprecia la generación simultánea de un resumen textual y un video-resumen a partir de los documentos seleccionados, lo que ilustra el funcionamiento del entorno como sistema de mediación cognitiva basado en fuentes delimitadas y trazables.
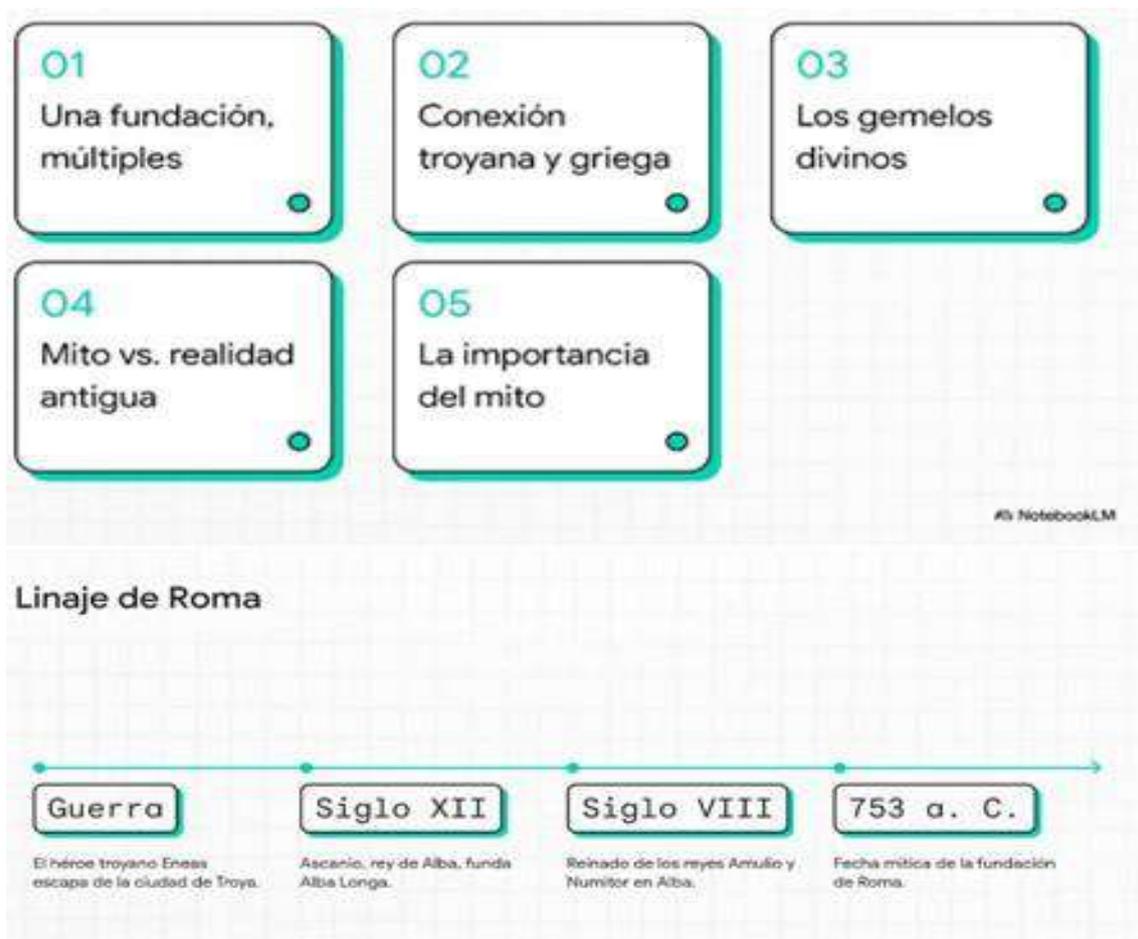
Figura 3. Capturas del video-resumen generado por *Notebook LM* a partir de los textos grecolatinos sobre la fundación de Roma. El material audiovisual ofrece una síntesis clara y accesible del relato fundacional, favoreciendo la comprensión global del proceso histórico; no obstante, tiende a armonizar las distintas tradiciones historiográficas, reduciendo la visibilidad de las divergencias ideológicas entre autores romanos y griegos.
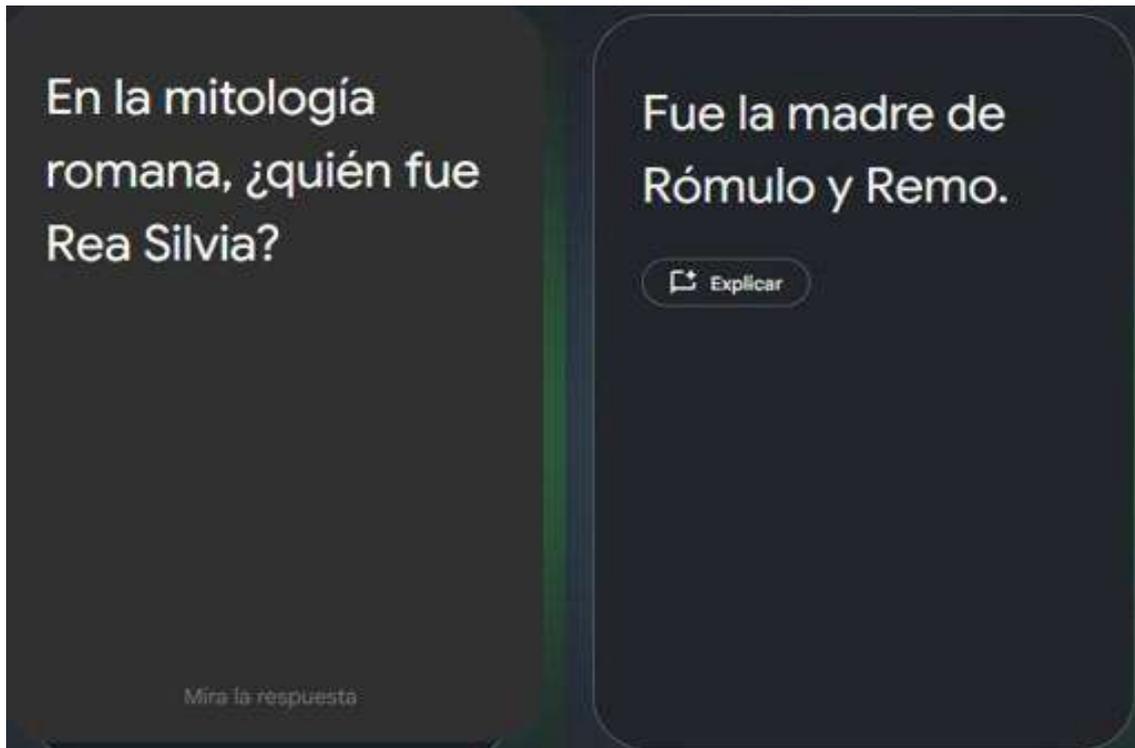
Figura 4. Ejemplo de tarjetas didácticas (*flashcards*) elaboradas por *Notebook LM* a partir del *corpus* documental. Este recurso permite reforzar conceptos clave, personajes y episodios mediante técnicas de repetición espaciada.

Let's *Chat* About Archaeology: Responsible and Thoughtful Use of AI Tools in Classroom, A Case Study

Yusi Liu

Bryn Mawr College
yliu2@brynmawr.edu
0009-0001-5719-1253

Daniel D'Elia

Bryn Mawr College
idelia@brynmawr.edu
0009-0009-1073-7276

Rocco Palermo

Bryn Mawr College
rpalermo@brynmawr.edu
0000-0002-0213-7276

ABSTRACT — This paper discusses an introductory archaeology class assignment designed to promote responsible and thoughtful use of AI in the classroom. We asked the students to take on two roles: first, as academic writers using AI tools, and second, as editors of writers' work. Our assignment encouraged students to think critically about how they frame their questions and how that framing influences the AI chatbot's responses. It also prompted them to evaluate the responses that AI gives them and raise their awareness of the legitimacy of the information AI provides. The fact-checking component also served an additional aim for archaeologists, as the quick spread of conspiracy theories online only increases the urgency for students to be trained in identifying historical and archaeological misinformation.

Overall, students' results demonstrate that this is a successful assignment. They paid close attention to the factual accuracy of the information, were observant of biased language in AI responses, and were able to critically evaluate issues with AI in academic learning and research. While this assignment can certainly be tweaked for greater effectiveness, particularly in relation to promoting better research skills, it served as a valuable exercise, not just for students to reflect on what constitutes ethical AI use in the classroom, but also for educators seeking a creative and engaging way to respond to the surge in the use of AI among students in general.

KEYWORDS — Archaeology, Pedagogy, Teaching, Generative AI, Critical Thinking.

1. INTRODUCTION AND RATIONALE

Since the release of ChatGPT in the fall of 2022, generative artificial intelligence (AI) has significantly grown in popularity in education, often marketed for its role in assisting users with gathering information, organizing thoughts, and generating ideas. Accordingly, discourses surrounding the use of AI in the classrooms have become especially pertinent as more students are gravitating towards AI tools for

assistance with assignments. While AI presents new opportunities for learning, it is also known to deliver incorrect or misleading information and undermine critical thinking, as one news headline laments, "Everyone Is Cheating Their Way Through College" (Walsh, 2025). How can university classrooms in the U.S. better prepare themselves for a time when they are welcoming a generation of incoming students whose literacy scores are at a historic low? When it comes to teaching Graeco-Roman antiquity, there is an added caution with the misuse and abuse of information, considering that various hate groups have historically and continuously exploited antiquity to promote nationalist and fascist propaganda.[1] Intertwined with this issue of the misappropriation of Graeco-Roman antiquity by white supremacist groups is the speed and ease with which pseudo-archaeology, typically heavily tied to white supremacist ideologies, gets propagated through the internet and other forms of digital media, such as podcasts, videos, social media, and online forums.[2] As Sean Rafferty observes, "Not only does the internet greatly magnify the availability of poor quality information, the interactive nature of the medium leads to the creation of 'echo chambers' that reinforce the perceived truth value of these ideas." (Rafferty, 2024, p. 18). To combat these historical and archaeological conspiracy theories in the classroom, Rafferty (2024, pp. 3-5) describes cultivating critical thinking skills among his students through teaching them how to evaluate the quality of sources and supporting evidence used within a given argument, in addition to the plausibility of the hypothesis's claims.

An assignment centered around critical thinking, therefore, not only serves to teach students how to evaluate responses derived from AI, but also how to identify this misinformation whenever and wherever they come across it, online or in the real world. Because, in the words of Heredia Chimeno (2025) in the inaugural volume of this journal, "AI is here to stay," it is more important than ever to encourage students to be mindful of their information sources and to critically evaluate the integration of AI in their academic journey. In this case study, we will share our experience with a class assignment we created to encourage responsible and thoughtful use of AI in classrooms. First, we will provide basic background information about our class and explain the motivation for designing this assignment for our syllabus. Then, we will highlight key observations from students' performance. Finally, we will discuss our conclusions and offer reflections for future AI-related teaching approaches.

---

[1] These appropriations are well documented by the platform *Pharos, Doing Justice to Classics*. https://pharos.vassarspaces.net/. (Accessed: 20 October 2025)

[2] On the very recent discussion pertaining to pseudo-archaeology, mainstream archaeology, and public archaeology, read Nathan J. Robinson's interview of Flint Dibble in "Why Joe Rogan Believes in Fake Archaeology" in *Current Affairs*. https://www.currentaffairs.org/news/flint-dibble. (Accessed: 29 October 2025)

2. THE ASSIGNMENT

With this backdrop of pervasive use of AI in higher education and the abovementioned concerns in mind, the authors of this paper taught the "Introduction to Classical Archaeology" (ARCH B.102) course at Bryn Mawr College in the spring semester of 2024–2025. Offered to students once every year, this course is a historical survey of the archaeology and art of Greece, Etruria, and Rome. It satisfies both the "Inquiry into the Past" and "Cross-Cultural Analysis" requirements for all bachelor's degrees at Bryn Mawr, and it also counts towards undergraduate major requirements in Archaeology, Classics, History of Art, and Museum Studies. Our class was at full capacity, with students from various backgrounds and at different stages of their educational progress. (Five first-year PhD students in Archaeology also attended the class as a requirement, though they had a different set of assignments and did not complete the assignment discussed in this case study.) The classroom thus included students ranging from those already pursuing an Archaeology degree to those in STEM fields taking the course to fulfill college-wide requirements, and from advanced students with proper experience in academic research and writing to freshmen with limited exposure to written assignments.

During syllabus planning, we discussed concerns about employing AI in the classroom and aspired to communicate its usefulness and potential drawbacks to students as part of the course's content. At Bryn Mawr College, we do not have an institutional policy on AI, and instructors are encouraged to design their own policy for their syllabi. A syllabus statement can inform students; it cannot, however, guarantee effectiveness in preventing them from using AI, nor in helping students to understand the complex issue of using AI in their academic work. Additionally, neither is it practical nor desirable to allocate time to lecture students about AI. Students at Bryn Mawr practice the Honor Code, which considers presenting or submitting AI-generated content a violation unless the instructor authorizes an assignment that allows it.[3] This may be more assuring than other universities without such an honor code system, as optimistically reported by Harper (2024), who considers that such infrastructure has already given students fewer incentives to cheat. Nevertheless, given the rise in students employing AI overall, we acknowledged that it is unavoidable to encounter a student using AI in at least some capacity, even if it does not cross the line into academic misconduct, such as generating topic ideas or providing grammar help. We deemed that AI can be useful, but it is important to have students understand the limitations of AI and what constitutes academic misconduct when it comes to AI, especially in an introductory course. In this sense, reiterating our Honor Code's emphasis on the use of AI when allowed by the instructor(s) can also offer a rather productive framework. By clearly

---

[3] The Honor Code is managed by Bryn Mawr's Self Government Association https://sga.blogs.brynmawr.edu/honor-board/honor-code/ (Access: 30 October 2025).

defining when and how AI tools can be used, we can motivate students to reflect on their own decision-making, evaluate the chatbot response, and critically develop a more knowledgeable approach to learning about the past.

The question then becomes: How do we translate these ongoing debates and concerns on using AI to students in the classroom? Our attempt at a solution was to design an assignment where students are encouraged to 1) think critically about how they frame their questions, and how it influences the way the AI chatbot responds, and 2) evaluate carefully the responses that AI gives them and raise their awareness to not just accept the information as true, but to identify areas where the AI gets it wrong or presents misleading information. In doing so, we intend to convey the message that they cannot always trust the responses that the AI gives and must do their own research. More importantly, by having students engage directly with the AI, we aimed to make the experience more interactive, encouraging an active learning process in which they could immediately observe the outcomes of their interactions. This approach was designed to reinforce and clarify the key concepts we wanted students to assimilate.

In the assignment "Ancient Past with Generative AI," we asked the students to take on two roles: first, as academic writers using AI tools, and second, as editors for the writers (see Appendix for our full assignment guidelines). As an academic writer, their task was to write a short scholarly encyclopedia entry on one of the ten themes the class had already covered (the Mycenaean world; the Minoan culture, the Iron Age and Lefkandi; the Greeks in Italy; the Persian Wars; the Peloponnesian Wars; the Athenian Acropolis; Greek symposium; Alexander the Great; and Roman presence in Greece). The students needed to pick an AI chatbot of their choice (such as ChatGPT or Google Gemini) and develop five questions for the AI to answer, each in 100–500 words, to help build this encyclopedia entry. We provided examples of encyclopedia entries and outlined their core elements, such as general history, sites and monuments, artistic developments, individuals, and objects. We asked students to be mindful of how they communicate with the chatbot to deliver their questions and receive the answers. Students were instructed to take at least five screenshots to document their interaction with the AI, which would show the questions they posed to the AI chatbot and its replies, while also proving that they completed the assignment. After the answers to their questions were generated, we then asked the students to imagine themselves in a more analytical capacity as editors working for the AI writer. Switching the role to editors, students were tasked with using citations and comments to explain why the information was correct, incorrect, or misleading. They were also asked to offer editorial suggestions on language organization and phrasing using the "Track Changes" mode. Finally, the students were requested to write a summary of 500–750 words explaining how they developed their questions for the AI, how they interacted with the AI, and how they went through the process of verifying and commenting on the AI responses. Students needed to include five screenshots to illustrate their summary.

3. RESULTS AND DISCUSSION

Overall, the students were attentive to checking the factuality of the information presented and noted instances where references were completely absent. In addition, students often pointed to areas where the AI could have provided additional details or examples in its responses. An analysis of the most frequently attested keywords within the assignment submissions found that the word mentioned most often by students was "accurate," being written a total of 71 times, with its related form "accuracy" appearing 17 times. The other six most repeatedly mentioned keywords are as follows: "detail," which was written 31 times, "scholarly" 35 times, "source" 28 times, "depth" 19 times, and "nuance" and "trust" 6 times. Less frequently used keywords include "verify" and "misinformation," which were used 5 times, closely followed by "bias," which appeared 4 times. Even rarer were keywords such as "overreliance," "inaccuracy," "verification," and "reliability," which all appear only once. The words "skeptical," "plagiarism," and "skepticism" do not feature at all within students' submissions (*Figure 1*). A keyword search on its own does not necessarily indicate much without contextual analysis, which is why we provide excerpts in the following paragraphs from submitted assignments, demonstrating how students have interpreted the guidelines and engaged with the core objectives of the assignment. In particular, as these keywords may suggest, common themes arose across the entire group, such as concern with the accuracy of the information, the quality of sources, the level of detail and nuance, and more general considerations regarding whether AI could be considered trustworthy.

With respect to the issue of accuracy and detail, one student, reflecting on the AI chatbot's answers to their questions on Mycenaean material culture, wrote, "Another issue I had discovered with the AI. . .was its use of blanket statements and generalizations. . .Though these statements were technically correct, having them without any explanation could easily become misleading." Another student remarked on the way AI cited information and made exaggerative statements, noting, "I don't like that AI can't directly cite sources, or that it makes broad claims that are sometimes difficult to debunk." Other critiques of the AI's responses also described them as repetitive and prone to generalization, with students often expressing a desire for greater nuance and variety, as one student commented, "I found that while it was somewhat factually accurate, it often made assumptions and did not delve deeply into many points." Another student, evaluating the utility of AI in the classroom, further stated that "in an academic context, I really don't think AI should have a place at the table. It's repetitive, unable to cite sources, vague, and at times makes concerningly broad generalizations. I can't see myself using this in the future or supporting anyone else in using it."

A smaller number of students noted the use of biased language in the presentation of information, observing, "In the answer to my second question, the

AI described Alexander's actions as 'remarkable'. . .To me, 'remarkable' feels like an opinionated word. Of course the AI cannot have an opinion, so I found this choice of language interesting." A handful of students also identified cases where the AI chatbot was presenting outdated and misleading information. One such case occurred in a student's assignment on Minoan culture, which found that the AI largely relied on scholarship by Arthur Evans and other twentieth-century scholars. The following comment by another student perhaps summarizes this issue well, "while AI can summarize historical narratives, it does not engage in the same critical discourse that historians and archaeologists do." Therefore, there is a concern that, even when AI does mention scholarly sources, it may present outdated narratives that have long been dismissed or challenged within present-day archaeological scholarship as settled fact, since it is incapable of critically evaluating scholarship on its own.

Furthermore, some students felt that the time and effort required to cautiously and extensively fact-check the AI enough to use the information it produced as research was inefficient, such as one student who wrote, "the user has no idea whether or not the software has directly copied the information word for word. I found myself triple checking the wording of my sources." While the word "plagiarism" is not mentioned, this student's discussion of how the AI can sometimes copy published work word for word is important to highlight here, as it indicates an awareness that AI can generate responses that are plagiarized. Another student also described having to conduct a significant amount of research to even begin fact-checking the AI, noting, "through this process of editing and fact-checking, I was forced to really research and learn about the history of the Acropolis in order to be able to effectively refute any mistakes."

Several students also reflected on the process of prompting the AI to produce an answer that was satisfactory in the first place. Multiple students mentioned having to rephrase their questions several times for the AI to generate a response that actually answered the questions they had in mind, and other students expressed frustration with the AI repeating material that it had already included in previous answers. For example, one student describes the process of reformulating their questions as follows: "At different points when I tried asking AI a question, it misunderstood or couldn't provide me with the answer I was looking for. This forced me to reevaluate my questions and adjust them to bring me to the results I was looking for." In this instance, the student describes not only how they noticed the AI limited the ways in which they could format their questions, but also how this adjustment step had to be repeated in different instances, interrupting their research process.

However, while students largely presented critiques of AI use, there were some instances where they could see AI serving a useful function, such as for "workshopping ideas," or "to save time, bridge intellectual dividends and process

data well," as two students wrote. One student also added, "I found AI great for summarizing information when I need a fast and digestible overview of a topic, which feels the same as googling something," which is important to consider as AI becomes increasingly integrated into search engines. In very few instances, a couple of students came away with or maintained the sense that AI could have fairly broad uses in academic and non-academic contexts, and they were not all too concerned by the current limitations of AI. It should be noted that, in these instances, these students missed key components of the assignment in their submissions, such as citations for their edits and screenshots of their conversation with the AI chatbot. Additionally, it should be mentioned that out of 33 responses, 2 were found to have used AI in the fact-checking editing procedure (we detected this by recognizing the cited publications were AI-fabricated) and in the composition of their summary, which undermined the purpose of the assignment and resulted in an automatic failure.

4. REFLECTION AND CONCLUSION

We consider this assignment a successful exercise that allows students to understand the concerns about AI applications in the classroom and academic assignments through a reflective lens. The majority of our students followed the assignment's instructions, and only a few missed some parts of the core requirement to include citations, screenshots, or a summary. This assignment created an opportunity for students to think more critically about AI. While most students (93.75 %, see *Figure 2*) reported no change in their negative perspective on AI use in academic contexts, it appears as though the majority were not inclined to use AI in the first place. While some students explicitly mentioned certain cases where AI use could be acceptable in an academic environment, most students overall indicated that they did not see AI as something that they would use themselves. It should be noted that we are not necessarily trying to convince students that they should never use AI—if they were to use it, they should know where AI is getting its information from and understand the biases that can exist in these sources, as well as the potential issues in how the AI presents information. From their summaries, most students were able to describe major pitfalls and ethical concerns that arise due to AI being used in academic research, indicating that, for most students, they had understood the aim of the assignment.

While we view this assignment as a successful effort to integrate AI into classroom teaching critically, there is considerable potential for future improvements. There are a few directions we would consider if we were to use this assignment in future classes. First, while many students critiqued the AI for being vague, repetitive, and generalizing, fewer students actually edited the phrasing and tone in their line edits of the AI's responses. To *prompt* students to think further about the phrasing and tone of AI's writing, we could state this expectation more explicitly in our guidelines and instruct students to discuss it in their summaries.

31

Secondly, we recognize that the assignment was less helpful in developing students' research skills. We noticed that, just as in the regular essay assignment, students tended to prioritize online sources over physical books. Only one student made a remark about the state of archaeological resources, noting that "a vast amount of archaeological data and historical archives still reside in physical libraries, museum collections, and excavation reports that remain undigitized." The majority of students used unreliable sources (such as Wikipedia, Britannica, and World History Encyclopedia) in their editorial process. When it comes to online academic sources, students struggle to tell the difference between outdated scholarship and newer arguments, and they often consider them equally valid and relevant. This tendency to focus solely on online academic articles is particularly concerning, as this information is integral to the AI ecosystem. For example, JSTOR, where most of our students find their scholarly articles, incorporated an AI research tool into its platform in July 2025.[4] Although there were only two cases where students used AI to generate what was supposed to be the non-AI parts of the assignment, this serves as a reminder of the importance of human intelligence in the research process and highlights the risks of false information and misuse of AI, where AI can easily produce fake evidence and publications at a large scale.[5] Additionally, these two incidents also prove that the Honor Code or a simple AI policy on the syllabus is not effective enough in preventing plagiarism with the use of AI in the classroom. For improvement, we suggest building more instruction on finding and evaluating academic sources into the class before this assignment.

Another issue was that many of the questions students asked the chatbot focused on basic facts rather than on interpreting archaeological evidence. This could be due to the assignment being given to a class of students who were mostly new to archaeology, considering some students were freshmen or had a background with limited research and writing experience. Taking this into account, the assignment might be more fitting for upper-level archaeology courses where students are already equipped with some level of archaeological methodologies and historical, factual knowledge to enable them to ask different questions and cite academic sources.

Beyond this goal of having students consider the limitations of AI and what constitutes ethical AI use, this assignment also sought to further develop research skills, to have students practice critical thinking, and to promote further engagement with the course material. As mentioned above, the fact-checking component of the assignment compelled students to actually undertake the research process themselves and to see how AI formulates errors or misleading answers in real-time. Through the process of editing and then summarizing their experience of

---

[4] https://www.about.jstor.org/blog/a-new-chapter-for-jstors-ai-research-tool/. (Access: 27 October 2025).
[5] See Haider *et al.* (2024) and Walters & Wilder (2023) on the discussion of ChatGPT-generated false publications.

both chatting with the AI and verifying the information provided to them, we found that we mostly succeeded in prompting reflection on ethical AI use, while still incorporating the larger aims of the course.

The fact-checking component also served an additional aim; for archaeologists, the quick spread of conspiracy theories online only increases the urgency for students to be trained in identifying historical and archaeological misinformation, which poses verifiable danger and harm due to their frequent connection with white supremacist ideologies. However, the utility of assignments like the one we outlined here is not limited to archaeology alone. Rather, its potential applicability extends cross-disciplinarily, as changing the selected themes does not alter the structure or broader aims of the assignment at all, and most, if not all, disciplines in the sciences and humanities involve at least some form of academic writing.

For educators grappling with how to address AI use in the classroom, we recommend communicating clearly and early on with students regarding classroom or institutional rules about what constitutes ethical AI use, in addition to an explanation as to why ethical AI use is defined in this way. An assignment much like the one we implemented here may also be worthwhile to consider, especially for educators whose course plan includes a research component. Guidelines and additional details for our assignment can be found in the Appendix. However, should we give this assignment again, we would likely have the students jump to the editorial phase by assigning them an already AI-generated essay to annotate, which would have the benefit of reducing the environmental impact of AI use and possibly make students who did not like having to personally interact with the AI more comfortable working on the assignment. While this assignment can certainly be tweaked for greater effectiveness, particularly in relation to promoting better research skills, it served as a valuable exercise, not just for students to reflect on what constitutes ethical AI use in the classroom, but also for educators seeking a creative and engaging way to respond to the surge in the use of AI among students in general.

5. ACKNOWLEDGEMENT

BIBLIOGRAPHY

Haider, J., Kristofer, S., Björn, Ekström. and Malte, R. (2024). 'GPT-fabricated scientific papers on Google Scholar: Key features, spread, and implications for preempting evidence manipulation', *Harvard Kennedy School Misinformation Review*, *5*(5), 1-16. doi: 10.37016/mr-2020-156.
Harper, T. A. (2024) 'ChatGPT doesn't have to ruin college: The power of a robust honor code—and abundant institutional resources', *The Atlantic*, 23 October 23. Available at:

https://www.theatlantic.com/ideas/archive/2024/10/chatgpt-vs-university-honor-code/680336/ (Accessed: 23 October 2025).

Heredia Chimeno, C. (2025) 'Opening the conversation: Rethinking how we teach antiquity in the age of AI', *AI & Antiquity*, *1*(1), 4-14. doi: 10.64946/aiantiquity.v1i1.editorial

Rafferty, S. (2024). *Mythologizing the past: Archaeology, history, and ideology*. Abingdon: Routledge.

Walsh, J. D. (2025) 'Everyone is cheating their way through college: ChatGPT has unraveled the entire academic project', *New York Magazine*, 7 May. Available at: https://nymag.com/intelligencer/article/openai-chatgpt-ai-cheating-education-college-students-school.html (Accessed: 20 October 2025).

Walters, W. H. and Wilder, E. I. (2023) 'Fabrication and errors in the bibliographic citations generated by ChatGPT', *Scientific Report*, 13, article 14045. doi: 10.1038/s41598-023-41032-5.

FIGURES
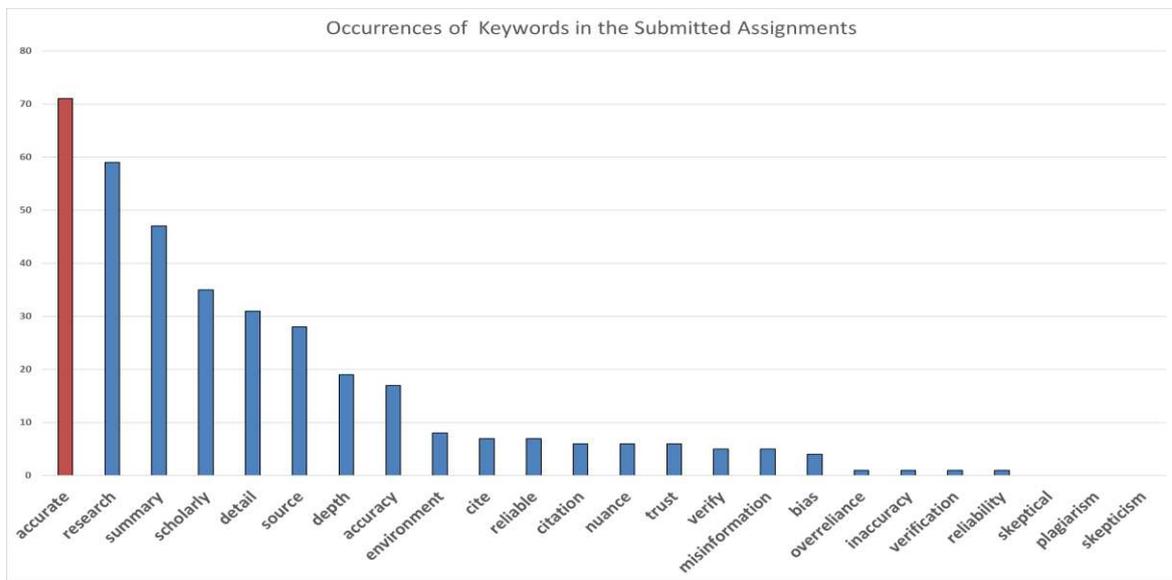
Figure 1. Chart demonstrating the most occurring keywords in the assignment submissions.
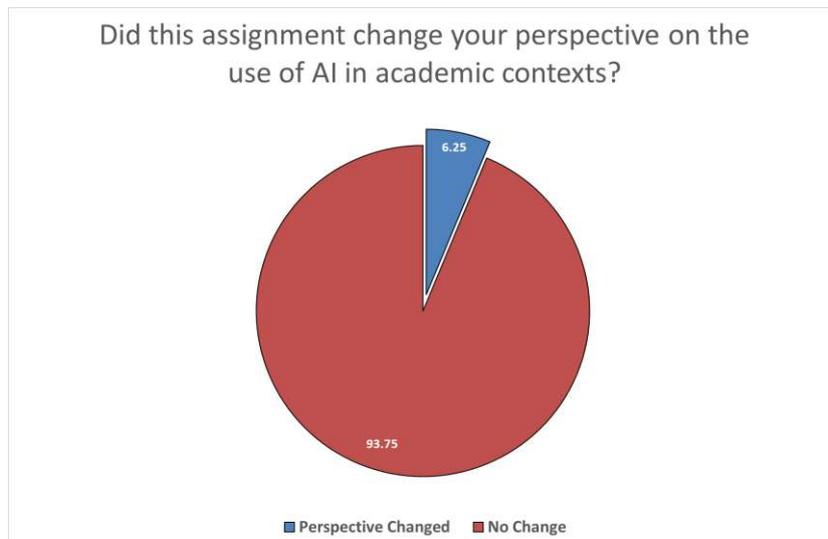
Figure 2. Percentage of students who reported a change in their perspective on the use of AI in academic contexts.

APPENDIX: GUIDELINES FOR THE ASSIGNMENT

The authors share the assignment guidelines with the broader community, with the intention that they be adopted, adapted, and improved for classroom teaching and research.

Artificial intelligence can assist us in gathering information, organizing thoughts, and generating and enhancing ideas. However, it can also produce false information that is difficult to detect, which can lead to plagiarism and undermine critical thinking. How do current generative AI tools perform when examining the ancient past? What is a more effective way to support your studies and research using these new tools? For this assignment, imagine you are an editor working for the AI writer. First, create five questions and ask an AI chatbot of your choice (such as ChatGPT, Google Gemini, etc.) to respond to each one (100-500 words each). Next, identify the accurate information the chatbot provides by offering citations and comments based on material gathered from class lectures and readings while also highlighting any incorrect or debatable sections with the appropriate citations. Finally, write a brief summary (about 500-750 words) explaining why and how you developed the five questions, the process of conversing with the bot, the editing and correction process, and your reflections on using generative AI to explore ancient histories. Additionally, include <u>five</u> screenshots to record your interaction with the chatbot.

Step-by-step Instruction:

1. Orient yourself to a specific theme. Choose a theme from below:
    o The Mycenaean world
    o The Minoan culture
    o The Iron Age and Lefkandi
    o The Greeks in Italy
    o The Persian Wars
    o The Peloponnesian Wars
    o The Athenian Acropolis
    o Greek symposium
    o Alexander the Great
    o Roman presence in Greece
2. Once you choose a theme, envision writing a short scholarly encyclopedia entry (an example of an entry authored by a student can be found on Moodle). When conducting inquiries, it is generally more effective to start with questions. What key facts do you need to create such an entry (for instance, chronology, geography, historical figures, battles, artifacts, etc.)? Additionally, what questions are essential to ask? Formulate five questions. Ensure these questions encompass history, sites/monuments, artistic developments, individuals, and objects.

3. Ask the generative AI of your choice to answer these questions. Be mindful of how you communicate with the chatbot to deliver your questions and receive the answers. Ensure that each response is between 100 and 500 words. *Don't forget to take screenshots during your interaction with the chatbot.*

4. After you receive the responses and finish your interaction with the chatbot, place them in a Word document (including your questions and conversations, not just the response).

5. Use either the Reviewing writing mode or a different font color to annotate the response. Identify which parts are correct and explain why. Highlight which parts are incorrect and provide reasons for that. How would you make corrections? Are there any sections that need clarification and further articulation? Would you phrase or rewrite any parts differently? Support your annotation with evidence from textbooks, class slides, or reliable scholarly sources.

6. After completing the annotation, write a brief summary of about 500 to 750 words that explains how and why you developed the five questions, your interaction process with the bot, the editing and correction steps you took, and your reflections on using generative AI to explore ancient histories. Be sure to include five screenshots to illustrate your summary.

7. Submit the AI-generated conversation with your annotations and your summary in a single document and upload it to Moodle.

Evaluating Generative AI in Historical Research: A Comparative Study on Identifying Primary Source Evidence in Ancient History

Raymond Solga

The College of Westchester
raymond.s.solga@gmail.com
 0000-0003-0689-0633

Mohamed J. Sarwar

The City University of New York
jahed.sarwar@gmail.com
 0009-0000-2405-4263

ABSTRACT — This study explores how traditional historical methods and generative AI tools compare in the identification, interpretation, and validation of primary sources in ancient history. Drawing from a dual case study approach—four case studies conducted by human historians and four by AI tools (GPT-4, Claude 2, Gemini, Perplexity)—we evaluate the epistemological strengths and limitations of each method. Using qualitative document analysis, historiographical criteria, and expert review, the study assesses source criticism, genre classification, provenance transparency, and evidentiary value. Results indicate that generative AI excels at broad content discovery and thematic synthesis but struggles with historical genre boundaries, source verification, and manuscript-based scholarship. Human researchers consistently outperform in contextual interpretation, critical chronology, and the adjudication of textual authority. We propose a human-in-the-loop framework combining digital speed with scholarly rigor, advocating for model pluralism, temporal prompting, and provenance-first protocols. This integrated methodology ensures AI contributes meaningfully to digital historiography without compromising historical standards.

KEYWORDS — Primary Sources, Ancient History, Historical Methodology, Generative AI, Digital Humanities

1. INTRODUCTION

The integration of generative artificial intelligence (AI) into the humanities has opened new pathways for research, synthesis, and discovery—particularly within the field of ancient history, where textual scarcity and interpretive complexity often pose major challenges. Large Language Models (LLMs) such as GPT-4, Claude, Gemini, and Perplexity now offer scholars rapid access to vast *corpora* of textual material, multilingual synthesis, and broad pattern recognition. However, these

efficiencies come with epistemological risks, including genre misclassification, hallucinated details, and opaque sourcing.

The reason Grok was not included is methodological rather than an oversight. At the time of data collection, Grok's access model, training transparency, and citation architecture differed substantially from the other tools examined (GPT-4, Claude 2, Gemini, and Perplexity). In particular, Grok's integration with real-time social media content and its evolving availability posed challenges for controlled, replicable scholarly comparison. To preserve methodological consistency and reproducibility, we therefore limited the study to models with more stable research-facing implementations. We note Grok as a relevant candidate for future studies as its scholarly affordances mature.

This paper presents a comparative evaluation of how AI tools and human historians perform when tasked with locating and interpreting primary sources across ancient historical topics. The methodology is bifurcated: one section assesses four historical inquiries answered using AI models; the other revisits the same research questions using traditional humanist methods. The aim is not to suggest one method replaces the other, but to measure their respective strengths and limitations using shared criteria—temporal framing, source provenance, linguistic fidelity, and historical reliability.

To structure this evaluation, we present eight case studies grouped into two categories. The first set of four—authored by Mohammed Sarwar—focuses on how generative AI models respond to queries about the Kaaba (Ibrahim and Ismail), Genghis Khan, Ibn Sina (Avicenna), and Jalaluddin Rumi. The second set—written by Raymond Solga—uses traditional historiographical methods to investigate the same kinds of questions across figures and topics such as Alexander the Great, the Hanging Gardens of Babylon, the Egyptian pyramids, and the Library of Alexandria.

By juxtaposing human-led and AI-assisted approaches, this article contributes to the growing field of digital epistemology and the ethics of human–AI collaboration in historical research. Ultimately, we argue that responsible integration of AI into historiography requires hybrid workflows grounded in scholarly methods—where digital tools act as accelerators of discovery, not arbiters of truth.

1.1 AI AS AUGMENTATION UNDER HUMAN EPISTEMIC SOVEREIGNTY

The article does not argue that AI replaces factual verification; rather, it demonstrates that AI augments scholarship by accelerating the pre-verification stages of research, not by resolving epistemic authority. This distinction is essential. Generative AI contributes value in three specific, bounded ways:

1. Rapid exploratory mapping of a research domain.
2. Surface-level clustering of names, texts, periods, and traditions that may otherwise remain siloed.

3. Cross-cultural and multilingual exposure, particularly across historiographies unfamiliar to the researcher.

These functions occur before verification and interpretation. In other words, AI operates as a cognitive accelerator, not as an epistemic judge. The article's findings show that even when AI responses are only marginally reliable in isolation, they still reduce the initial cost of exploration—the time and labor required to identify what might need to be verified. That reduction is the augmentation.

The exercises work because the authors already know how to verify. This observation is correct—and intentional. The study is not designed to test whether AI can independently "do history." Instead, it evaluates how AI behaves when inserted into a workflow governed by expert verification norms. This mirrors how new research instruments are historically assessed:

4. Archaeological tools are validated by archaeologists.
5. Statistical software is validated by statisticians.
6. Digital archives are validated by trained historians.

AI, in this sense, is treated as methodologically immature instrumentation, not as an autonomous researcher. The authors' expertise is not a confounder; it is the control mechanism that allows epistemic failure modes (genre collapse, provenance loss, temporal drift) to be reliably observed and categorized.

This raises a legitimate question: if the authors already had the knowledge, what was actually learned from AI? What the authors learned was not new historical facts, but new structural insights into how AI mediates knowledge:

- Which types of sources are systematically surfaced or suppressed.
- How popularity and digital visibility distort historical authority.
- Where AI consistently conflates devotional, literary, and empirical genres.
- How different models reproduce distinct cultural and linguistic biases.

This is discovery at the methodological level, not the factual one. The article therefore contributes not new historical claims, but a diagnostic map of AI behavior—knowledge that historians, librarians, and educators did not previously have in a documented, comparative form. That diagnostic insight is itself a scholarly contribution.

But, where is the discovery facilitation? Discovery facilitation occurs in scope expansion, not validation. AI proved most useful in:

- Revealing adjacent traditions the researcher might not initially query.
- Surfacing parallel corpora across Islamic, Persian, Mongol, Chinese, and Greco-Roman historiographies.

- Exposing how the same historical figure is framed differently across cultures and centuries.

For example, while AI failed to authenticate manuscripts for Ibn Sina or Rumi, it consistently surfaced interdisciplinary linkages (medical, philosophical, mystical, literary) that informed subsequent human-guided inquiry. These linkages represent research leads, not conclusions. Thus, AI facilitates discovery of questions, not answers.

The hard question: if scholars use AI to research what they don't already know—who verifies? Verification cannot be delegated to AI. It must reside in one of three human-centered structures:

- Expert scholars, in advanced research contexts.
- Instructional scaffolding, in student and training environments.
- Institutional validation systems, such as libraries, archives, and peer review.

1.2 METHODOLOGICAL FRAMEWORK FOR RESPONSIBLE AI INTEGRATION

The article's four-pillar framework (temporal prompting, provenance-first methodology, model pluralism, and human-in-the-loop validation) is designed precisely to answer this concern. It provides a way for non-experts to use AI responsibly by embedding verification responsibility where it already exists institutionally.

In pedagogical contexts, this means:

- Students use AI to generate hypotheses and leads.
- Verification is performed through curated sources, librarians, faculty, or archival tools.
- AI output is treated as provisional and auditable, not authoritative.

In research contexts, this means:

- AI accelerates exploratory breadth.
- Human scholars retain adjudicative sovereignty.

The article does not claim that AI produces reliable historical knowledge on its own. It claims that AI can responsibly augment scholarship only when epistemic authority remains human. In that sense, AI's role is analogous to:

- A research assistant who gathers materials but does not interpret them.
- A discovery layer that expands visibility without conferring legitimacy.

This framing preserves scholarly rigor while acknowledging genuine technological utility. This study examines the role of generative artificial intelligence in ancient historical research by comparing traditional human-directed

historiographical methods with AI-generated research outputs evaluated under human supervision. The aim is not to assess whether AI can replace historical expertise, but to determine how, where, and under what constraints AI systems may responsibly augment scholarly research practices in Ancient Studies.

To achieve this aim, the article employs eight case studies, divided evenly between two methodological approaches. Four case studies apply traditional historical methods grounded in source criticism, provenance analysis, temporal framing, and genre differentiation. Four corresponding case studies examine how generative AI systems respond to comparable historical research questions when prompted under controlled conditions. This paired design allows for direct methodological comparison and highlights both the strengths and limitations of AI-assisted research relative to established historiographical standards.

The study further advances a four-pillar framework for responsible AI integration—temporal prompting, provenance-first methodology, model pluralism, and human-in-the-loop validation—intended to support inclusive, pedagogically transformative approaches to Ancient Studies. In doing so, the article aligns with the mission of AI & Antiquity by rethinking how digital tools can be ethically and effectively integrated into historical research and teaching without compromising disciplinary rigor.

2. TRADITIONAL HISTORICAL METHODOLOGIES

What counts as a primary source in Ancient History? In the discipline of ancient history, the term "primary source" holds a precise and contextual meaning. It refers to materials contemporaneous with the events or periods under study—inscriptions, coins, papyri, cuneiform tablets, chronicles, and other forms of direct evidence produced by historical actors themselves. These sources are distinguished by their proximity in time and space to the subjects they describe, often preserved through manuscripts, artifacts, or oral traditions subsequently recorded.

However, unlike modern archival history, ancient historiography frequently deals with textual fragments, later redactions, and source compilations created long after the events they describe. As such, human historians apply a layered interpretive lens—treating even canonical works (like Herodotus' *Histories* or Ibn Ishaq's *Sīrah*) as potentially derivative rather than strictly "primary," unless anchored by corroborating archaeological or epigraphic evidence.

This granularity is essential for distinguishing direct eyewitness accounts from later literary, theological, or philosophical renderings, a distinction often blurred in generative AI outputs. Indeed, human scholars rely on provenance-based evaluation—a methodological discipline rooted in identifying the physical, editorial, and contextual origin of sources. For instance, the Canon of Medicine by Ibn Sina is not evaluated merely by its contents, but also by its transmission history, the

manuscript shelf marks in European and Islamic libraries, and its translation lineage across centuries and regions.

Temporal framing serves to locate a source not only within a timeline but within a historiographical debate: was this text produced by a participant in the events, a compiler of older traditions, or a commentator removed by centuries? Textual criticism further dissects interpolations, scribal modifications, and regional editions, allowing historians to treat the "text" not as fixed, but as a process—understanding what version was read, when, by whom, and how it was interpreted. Such rigor cannot be easily approximated by AI models, which often treat all references as equally valid unless instructed otherwise.

Raymond Solga's analysis of ancient figures such as Alexander the Great, the Pyramids, and the Library of Alexandria demonstrates how interpretive judgment is deployed to filter, contextualize, and synthesize scattered historical fragments into coherent narratives. These interpretations hinge not only on access to sources but on training in genre recognition, regional historiography, and comparative critique. Mohammed Sarwar's counterpart analysis of AI behavior shows how these same interpretive tasks break down in digital contexts. Generative models, despite their fluency, lack the ability to distinguish devotional narratives from empirical accounts (as seen in the Kaaba case) or to critically evaluate manuscript chains (as with Avicenna).

Together, both lenses reveal a crucial insight: historical reasoning is not only about information retrieval, but about epistemological curation—knowing which sources to trust, why they matter, and how they fit within larger systems of meaning.

3. HUMAN-DIRECTED CASE STUDIES (TRADITIONAL HISTORIOGRAPHICAL METHOD).

The four case studies presented in section 3 were conducted using traditional historical research methodologies and were authored entirely by the human researchers. No generative AI tools were used in the identification, interpretation, or validation of primary sources in these analyses.

Each case study applies established historiographical practices, including provenance-based source evaluation, chronological contextualization, genre differentiation, and critical interpretation. These analyses serve as a methodological baseline, illustrating how expert human scholars approach ancient historical evidence when operating within conventional academic frameworks.

The purpose of this section is to establish a clear reference point against which AI-generated research behaviors can be meaningfully evaluated. The analyses reflect disciplinary expertise and interpretive judgment rather than algorithmic synthesis, ensuring that subsequent comparisons with AI-generated outputs are methodologically grounded and analytically coherent.

3.1 CASE STUDY: ALEXANDER THE GREAT AND THE HISTORIOGRAPHY OF EMPIRE

Research Focus

- *Prompt*: What primary sources document the campaigns and leadership of Alexander the Great, and how do historians assess their credibility?
- Primary Sources Cited:
    - Arrian, *Anabasis of Alexander* (2nd century CE)
    - Plutarch, *Life of Alexander* (1st–2nd century CE)
    - Diodorus Siculus, *Bibliotheca Historica* (1st century BCE)
    - Quintus Curtius Rufus, *Histories of Alexander the Great* (1st century CE).
    - Justin (Marcus Junianus Justinus), *Epitome of the Philippic History of Pompeius Trogus* (2nd–4th century CE)

    These sources are all post-Alexandrian, compiled by Roman and Hellenistic authors based on now-lost contemporary materials.

Assessment and Analysis

Classical accounts of Alexander's campaigns were written well after his death (323 BCE), often shaped by political agendas or moralistic storytelling traditions. Arrian is generally viewed as the most reliable due to his critical engagement with earlier sources such as Ptolemy and Aristobulus, both of whom were contemporaries of Alexander. However, their original texts are lost, and Arrian's reliance is filtered through his own historiographical lens.

Plutarch and Diodorus Siculus present Alexander in more literary and moralized terms, emphasizing themes of virtue, ambition, and tragedy rather than strict military or geopolitical analysis. Curtius Rufus, although engaging, is stylistically dramatic and lacks consistent chronology. Justin offers a compressed and often unreliable narrative. Despite their limitations, these texts form the backbone of Alexander scholarship, interpreted through cross-referencing, archaeological context, and philological reconstruction.

Interpretation

The traditional historian must perform careful triangulation: cross-analyzing ancient texts with material remains (coinage, inscriptions, architectural evidence) and evaluating the biases of each author. For instance, while Arrian presents Alexander as a rational and disciplined leader, other accounts suggest a descent into tyranny and hubris. The diversity of portrayals underscores the contested legacy of Alexander as both a liberator and a conqueror.

Importantly, the human researcher distinguishes between:

- Primary proximity (authors who used eyewitness testimonies),

- Genre (moral biography vs military logbook),
- Transmission chain (lost originals vs extant summaries).

This approach exemplifies the provenance-first historical method, where textual authority is built not just on content but on chronology, context, and credibility (see *Table 1*).

3.2 CASE STUDY: THE HANGING GARDENS OF BABYLON – MYTH, MEMORY, AND MATERIAL EVIDENCE

<u>Research Focus</u>

- *Prompt*: What primary evidence supports the existence of the Hanging Gardens of Babylon, and how have historians assessed the credibility of these sources?
- Primary Sources Cited:
    - Berossus, *Babyloniaca* (3rd century BCE; fragmentary, via later citations).
    - Strabo, *Geography* (1st century BCE).
    - Diodorus Siculus, *Bibliotheca Historica* (1st century BCE).
    - Quintus Curtius Rufus, *Histories of Alexander* (1st century CE).
    - Philo of Byzantium, *De septem orbis spectaculis* (disputed authorship; possible 3rd century BCE–2nd century CE).

While the Hanging Gardens are considered one of the Seven Wonders of the Ancient World, no definitive Babylonian cuneiform records or archaeological structures have been directly identified with them in Babylon.

<u>Assessment and Analysis</u>

Classical authors such as Diodorus, Strabo, and Curtius describe an elaborate, multi-tiered garden built by Nebuchadnezzar II or, according to some sources, an Assyrian ruler like Sennacherib. These narratives appear centuries after the gardens were supposedly constructed and rely on secondhand or legendary material.

Notably, Berossus—a Babylonian priest writing in Greek—linked the gardens to Nebuchadnezzar, but his writings survive only through citations in later Roman sources. Some modern archaeologists suggest that confusion with Assyrian palatial gardens in Nineveh, which had complex irrigation systems, may have led to the myth being transferred to Babylon. Recent scholarship, including satellite archaeology and reinterpretation of Assyrian inscriptions, supports this theory (Dalley, 2013). Thus, the material culture evidence diverges from literary tradition.

<u>Interpretation</u>

This case reveals the disjunction between literary memory and archaeological visibility. While the Hanging Gardens occupy a central place in classical imagination,

there is no definitive Babylonian or Mesopotamian documentation for their existence in Babylon itself.

Historians face a dual challenge:

- Decoding the rhetorical intent of classical authors (wonder literature, exoticism).
- Tracing misattributions or cultural conflations across empires and centuries.

The historiographical debate emphasizes the importance of cross-disciplinary triangulation—philology, archaeology, and comparative literature—to adjudicate the credibility of ancient "wonders" (see *Table 2*).

3.3 CASE STUDY: THE PYRAMIDS AND EGYPTIAN MONUMENTALISM – ALIGNING ARCHITECTURE WITH POWER

Research Focus

- *Prompt*: What primary sources support the construction history of the Egyptian pyramids, and how do archaeological findings align with dynastic records?
- Primary Sources Cited:
    - Pyramid Texts (Old Kingdom funerary inscriptions, c. 2400–2300 BCE).
    - Palermo Stone (Royal Annals of the Old Kingdom, fragmentary).
    - Herodotus, *Histories* (5th century BCE).
    - Diodorus Siculus, *Bibliotheca Historica* (1st century BCE)
    - Inscriptions at Wadi el-Jarf (logistical records related to Khufu's pyramid).
    - Modern archaeology adds to this with material sources:
        - Worker's graffiti in Khufu's pyramid complex
        - Quarrying sites (e.g., Tura, Aswan)
        - Logistics *papyri* (e.g., the "Diary of Merer")

Assessment and Analysis

The Egyptian pyramids—particularly the Great Pyramid of Giza—are among the best documented ancient monuments due to their monumental scale and the durability of stone inscriptions. However, a split exists between:

- Contemporary Egyptian administrative records, and
- Later Greco-Roman interpretations that often introduce fantastical or moralized accounts.

Herodotus, for instance, wrongly claimed the pyramids were built by enslaved masses, a view long propagated in popular imagination. In contrast, the Wadi el-Jarf

papyri (c. 2600 BCE), discovered in 2013, include detailed logbooks from royal overseers, offering direct insights into daily logistics of pyramid construction.

These findings support the thesis that the pyramids were built by a corvée labor force of skilled Egyptian workers rather than slaves, embedded in a religious and political ideology that linked pharaohs with divine order (Ma'at).

Interpretation

The pyramid case exemplifies harmonization between material evidence and indigenous textual traditions. While classical authors provide external narratives, it is the Egyptian epigraphy—pyramid texts, administrative records, graffiti—that ground the architecture in a historical-ritual framework.

Unlike the Kaaba or Hanging Gardens, the pyramids benefit from contemporaneous primary sources, including royal annals and internal inscriptions. This allows researchers to:

- Align symbolic meaning with historical context,
- Assess labor organization through recovered papyri, and
- Reconstruct inter-regional logistics and resource flows.

It also highlights the limitations of classical ethnography (like Herodotus) in reconstructing ancient history without access to local archives or scripts (see *Table 3*).

3.4 CASE STUDY: THE LIBRARY OF ALEXANDRIA AND INTELLECTUAL MEMORY

Research Focus

- *Prompt*: What are the primary sources confirming the existence and destruction of the Library of Alexandria, and how credible are these accounts?
- Primary Sources Cited:
    - Strabo, *Geography* (1st century BCE – 1st century CE).
    - Plutarch, *Life of Caesar* (1st century CE).
    - Aulus Gellius, *Attic Nights* (2nd century CE).
    - Ammianus Marcellinus, *Res Gestae* (4th century CE).
    - Letter of Aristeas (pseudepigraphal, 2nd century BCE).
- Secondary interpretations and late commentaries:
    - Edward Gibbon, *The Decline and Fall of the Roman Empire* (18th century).
    - Carl Sagan, *Cosmos* (20th century popularization).

Assessment and Analysis

Unlike monumental architecture, the Library of Alexandria leaves behind no archaeological footprint—its memory survives solely through textual allusions. The absence of direct material remains, or institutional records creates interpretive ambiguity.

Accounts such as Plutarch's suggest that the library was accidentally burned during Julius Caesar's Alexandrian campaign (48 BCE). Others, like Ammianus, imply it persisted into Late Antiquity, while Arabic sources claim it was destroyed under Caliph Umar—though this is widely considered apocryphal by modern historians.

No contemporary document confirms the foundation or cataloguing of the library, making reconstructions speculative. The most accepted view is that the library functioned under the Ptolemies (3rd–2nd centuries BCE), possibly as part of the Mouseion (House of Muses)—a scholarly complex rather than a standalone building.

Interpretation

The Library of Alexandria functions more as a cultural symbol than a documented historical institution. Its imagined destruction—whether by Caesar, Christians, or Muslims—reflects intellectual anxieties about knowledge loss more than historical fact.

This case is unique among the four because:

- The primary "evidence" is largely retrospective and rhetorical, not archival.
- There is no archaeological record of the building.
- Accounts are shaped by ideological bias—e.g., Enlightenment critiques of religious fanaticism or 20th-century appeals to scientific humanism.

Despite this, the library's legacy remains foundational to discourses on information preservation, scholarly authority, and the vulnerability of cultural memory. As a historiographic construct, it reveals how absence of evidence becomes a canvas for projection—often more revealing of the source's context than of the event itself (see *Table 4*).

Transitional Rationale: From Human Interpretation to AI-Generated Analysis

The human-directed case studies presented above demonstrate how historical knowledge is constructed through expert judgment, evidentiary scrutiny, and contextual reasoning. These studies establish the analytical criteria—source reliability, provenance transparency, temporal accuracy, and genre awareness—that guide responsible historical scholarship.

The following section applies these same criteria to AI-generated research outputs addressing the same historical questions and evidentiary domains. By holding AI systems to identical evaluative standards, the article enables direct comparison between human and AI-mediated research practices, clarifying where generative tools align with, diverge from, or fall short of traditional historiographical norms.

4. EVALUATING GENERATIVE AI IN HISTORICAL RESEARCH

The preceding sections established the epistemological foundation of historical inquiry—anchored in temporal framing, provenance, and interpretive rigor—and demonstrated its application through four human-guided case studies: Alexander the Great, the Hanging Gardens of Babylon, the Pyramids, and the Library of Alexandria. Building upon these benchmarks, section 4 evaluates the performance of generative artificial intelligence (AI) systems when assigned comparable historical research tasks. The objective is to determine whether these systems can approximate scholarly reasoning or if they merely replicate surface-level information. This evaluative phase functions as a bridge between classical historiography and digital epistemology, reflecting how technological tools may expand access while still demanding critical human oversight.

Prompt Design and AI Behavior

Generative AI models operate through probabilistic language prediction; consequently, the integrity of their historical outputs depends on prompt architecture. In a scholarly context, *prompt design* must emulate the methodological discipline human historians apply when interrogating sources. Effective prompts include the following elements:

- Temporal specificity – Define the chronological range (e.g., "sources produced between 1200 and 1250 CE") to mitigate anachronistic conflation.
- Provenance-first framing – Require shelf marks, codices, or archival repositories rather than generic web references.
- Genre delineation – Separate empirical documentation from devotional or literary interpretation.
- Model pluralism – Submit identical prompts across multiple AI systems (GPT-4, Claude 2, Gemini, Perplexity) to expose variance and bias.
- Human-in-the-loop review – Treat AI outputs as leads for discovery, not as authenticated evidence.

When these controls are absent, AI models tend to conflate theological or poetic traditions with empirical historiography, prioritize popularity over authority, and omit critical metadata such as manuscript lineage or edition provenance. These behavioral trends are examined in greater depth within the AI case studies that follow.

Method: Mixed-Model Prompting and Response Evaluation

A *mixed-model prompting design* was adopted to compare AI behaviors under standardized and variable conditions. Each model—GPT-4, Claude 2, Gemini, and Perplexity—received identical baseline prompts later refined with temporal and genre filters. Outputs were assessed using five historiographical criteria drawn from traditional research methodology:

- Temporal proximity – Accuracy in situating a source within its historical timeframe.
- Provenance transparency – Presence of verifiable editions, repositories, or catalog numbers.
- Genre correctness – Ability to distinguish theological, literary, and empirical content.
- Linguistic fidelity – Use of original titles and respect for translation lineage.
- Evidentiary reliability – Alignment with accepted historiographical standards.

This evaluative procedure mirrors the interpretive rigor used by human historians, transforming scholarly heuristics into a digital analytical rubric. It maintains continuity with the paper's earlier sections while translating epistemological discipline into computational assessment.

Comparison with a Traditional Search Engine

Our study intentionally focused on comparative epistemic behavior among generative AI systems, rather than on retrieval efficiency alone. Traditional search engines primarily function as indexing and ranking tools, leaving source evaluation and synthesis entirely to the researcher. By contrast, generative AI systems act as interpretive intermediaries, producing synthesized narratives that may obscure provenance and genre boundaries.

That said, the manuscript now clarifies this distinction more explicitly by framing search engines as a baseline discovery layer, against which generative AI represents a qualitatively different epistemic intervention. We agree that future research could profitably include a controlled search-engine comparison to further isolate where AI meaningfully departs from established research workflows. Also, there is a summary of models tested (see *Table 5*).

Digital Epistemology and AI in Historical Research

Within the broader discourse of *digital epistemology*, generative AI functions as a novel *discovery layer* in historiography (Graham, 2020; Small and Green, 2021). These systems demonstrate computational breadth—rapid information retrieval, multilingual synthesis, and thematic clustering—but lack interpretive depth. Their

probabilistic reasoning favors semantic plausibility over evidentiary verification, producing results that may appear scholarly yet remain epistemically opaque.

As Sarwar (2025) observes, AI acquires analytical value only when embedded within human-supervised workflows that secure provenance, contextual integrity, and chronological control. In such a configuration, technology accelerates access and pattern detection, while scholars retain adjudicative authority over meaning and validity. Responsible integration therefore rests on four methodological pillars—temporal prompting, provenance-first reasoning, model pluralism, and human-in-the-loop validation—which structure the comparative framework developed below.

Section 4 functions as the methodological bridge between classical historiography and computational research. By translating humanistic criteria into prompt architecture and model assessment protocols, it reframes discovery and verification not as competing phases but as interdependent operations within digital scholarship. The following case studies operationalize these criteria, identifying the conditions under which algorithmic synthesis extends historical insight and the points at which epistemic limits require renewed human interpretation.

The framework formalized here thus proposes a responsible model for AI-assisted historical research—one in which computational procedures enhance exploratory capacity while the interpretive act remains irreducibly human.

5. AI CASE STUDIES IN SOURCE DISCOVERY

Section 5 examines research outputs generated by four generative AI systems (GPT-4, Claude 2, Gemini, and Perplexity) when tasked with historical research prompts corresponding directly to the human-directed case studies in section 3. All *prompts* were issued to each model under identical conditions, enabling comparative evaluation across platforms.

References to individual models within subsection headings are illustrative rather than exclusive, highlighting characteristic response patterns rather than isolating model-specific testing. The methodological objective is to assess systemic behaviors across generative AI systems, not to privilege or single out any one platform.

AI-generated outputs are treated here as objects of analysis, not as authoritative historical narratives. The role of the human researcher is evaluative and interpretive: assessing how AI systems identify sources, handle provenance, differentiate genres, and manage temporal framing when confronted with complex ancient historical questions.

Section 5 operationalizes the evaluative framework established in section 4 by applying the mixed-model prompting method across four thematic case studies: the Kaaba, Genghis Khan, Ibn Sina (Avicenna), and Jalaluddin Rumi. Each case probes how large-language models (LLMs) handle questions of provenance, chronology, and genre when identifying primary sources within culturally and temporally complex historical domains. The findings illustrate the epistemological strengths and weaknesses of generative AI and reinforce the necessity of human interpretive oversight.

5.1 CASE STUDY: THE KAABA AND THE TRADITION OF IBRAHIM AND ISMAIL

Prompt to GPT-4: *"List primary sources documenting the construction of the Kaaba by Ibrahim and Ismail, with historical citations".*

<u>Model Behavior</u>

GPT-4 returns references to Quranic verses (e.g. 2: 125–127), biographies from Ibn Ishaq and al-Tabari, and Islamic summaries from online encyclopedias. No contemporaneous archaeological record or inscription is returned for the time of its purported construction.

<u>Analysis and Interpretation</u>

The Qur'an and classical tafsir (commentary) are foundational for the Islamic tradition, but they do not constitute contemporaneous sources in a historiographical sense. Ibn Ishaq and al-Tabari, centuries later, provide narratives rooted in oral tradition. GPT-4 fails to differentiate religious testimony from empirical historical sources. This is one case in which LLMs cannot be expected to do this without highly precise, scholarly-level prompting and an understanding of genre or temporal framing (see *Table 6*).

In this case study, the four models (GPT-4, Claude 2, Gemini, and Perplexity) converged on the conventional Islamic historical narrative, relying predominantly on scriptural and exegetical traditions. GPT-4 referenced both the Qurʾān and Ibn Isḥāq's *Sīrah*, yet did not distinguish between theological transmission and primary source analysis as defined within historical methodology. Claude 2 produced a strictly confessional account, offering no assessment of provenance, textual stratification, or historiographical reliability. Gemini expanded the narrative with additional summary material, but without verifiable citations, suggesting dependence on secondary popular platforms such as IslamicFinder or general encyclopedic repositories.

Perplexity further reduced the response's scholarly utility by combining modern commentary with aggregated web content, thereby blurring the boundary between historical evidence and contemporary interpretation. Across all four outputs, there was a consistent absence of engagement with archaeological data,

epigraphic corpora, or critical historiographical research. Consequently, while the models reproduced a coherent traditional narrative, they failed to meet the evidentiary standards required for source-critical historical inquiry.

5.2 CASE STUDY: GENGHIS KHAN AND MONGOL HISTORIOGRAPHY

Prompt to Claude: *"Identify primary sources about Genghis Khan's campaigns, with date of composition and authorship".*

<u>Model Behavior</u>

Claude provides sources including The Secret History of the Mongols, Juvayni's Tarikh-i Jahan Gusha, and Chinese court histories like the Yuan Shi. The model establishes that The Secret History belongs to the 13th century and recognizes its anonymous authorship.

<u>Analysis and Interpretation</u>

Although Claude demonstrated greater surface accuracy than GPT-4's earlier response, it omitted key historiographical parameters, including the *Secret History*'s extra-court authorship and the delayed transmission of relevant Persian and Chinese textual traditions. Moreover, the analysis did not engage with manuscript traditions, critical editions, or institutional repositories capable of establishing documentary provenance (see *Table 7*).

When prompted to disclose sources on Genghis Khan, all examined models conflated Mongol, Persian, and Chinese materials without constructing a chronological framework or identifying the manuscript lineages underpinning these traditions. GPT-4 correctly named the *Secret History of the Mongols* and Rashīd al-Dīn's *Jāmiʿ al-tawārīkh*, yet characterized the former as an "official history" and failed to situate either text within the temporal and historiographical distance required for source-critical evaluation. Claude 2 and Gemini reproduced the same source set in generalized form, offering narrative summaries without citation granularity or methodological assessment. Perplexity, by contrast, produced aggregated, blog-like explanations detached from identifiable primary or scholarly editions.

Across all systems, there was a persistent inability to distinguish between firsthand chronicles, later compilations, and modern secondary interpretations. This pattern exemplifies a broader problem of flattened historicity, whereby AI models collapse multi-layered textual traditions into a single informational stratum, supplying content without the analytical apparatus necessary to evaluate provenance, transmission, and evidentiary hierarchy.

5.3 CASE STUDY: IBN SINA (AVICENNA) AND SCIENTIFIC MANUSCRIPT TRADITIONS

Prompt to Gemini: *"List Ibn Sina's original works in Arabic, their composition dates, and surviving manuscripts".*

<u>Model Behavior</u>

Gemini produced a generic list: Three key historical texts include The Canon of Medicine, The Book of Healing along with Remarks and Admonitions. The reference list relied on Western-language summaries as well as internet encyclopedias rather than using manuscript catalogs or critical Arabic editions.

<u>Analysis and Interpretation</u>

One of the common challenges faced by AI-supported humanities research is its tendency to present primary and tertiary sources as one unified list of results. Gemini provides references without mentioning original Arabic collections such as al-Azhar or Topkapi, nor even to reputable critical editions (e.g. the Avicenna Latinus project). The absence of a human reviewer to verify sources leads to an isolated study that lacks proper context (see *Table 8*).

In exploring Ibn Sina's (Avicenna's) contributions, particularly *The Canon of Medicine*, the tools exhibited another pattern of epistemic slippage. GPT-4 accurately named key texts but referenced them mostly through Latin or English editions, ignoring Arabic originals or manuscript repositories like those at Istanbul's Süleymaniye Library. Claude 2 offered Western academic summaries and biographical sketches, often detached from textual transmission history. Gemini repeated this pattern, citing Ibn Sina through encyclopedic entries without identifying edition lineage or commentary traditions. Perplexity added to the confusion by referencing blog posts and simplified web articles. None of the AI systems included citation anchors to critical editions, manuscript IDs, or catalog references. This reveals a systemic gap in AI's handling of scientific-historical content: while it can summarize philosophical or medical achievements, it fails to engage with the documentary infrastructure—codices, scripts, translation chains—that underpin Islamic medical historiography.

5.4 CASE STUDY: JALALUDDIN RUMI AND PERSIAN SUFI LITERATURE

Prompt to GPT-4: *"Cite Persian primary editions of Rumi's Masnavi with manuscript locations or catalog references".*

<u>Model Behavior</u>

GPT-4 offers summaries of Rumi's philosophy and refers to Nicholson's English translation of the *Masnavi* as a "primary text." It fails to cite Persian manuscript repositories, such as those in Konya, Tehran, or the Süleymaniye Library in Istanbul.

Analysis and Interpretation

Rumi's *Masnavi*, one of the most influential works of Sufi literature, survives in multiple early Persian manuscripts. Citing modern English translations as "primary" misrepresents both the text's origin and its interpretive tradition. Once again, the AI system demonstrates fluency in narrative but lacks bibliographic literacy or linguistic fidelity (see *Table 9*).

The examination of Rumi's historical presence through AI output underscores another form of epistemological distortion. GPT-4 frequently returned verses from the *Masnavi* but cited them via modern poetic renderings (e.g., Coleman Barks), lacking any reference to the original Persian texts or critical editions like those curated in Konya or Tehran. Claude 2 presented mystical interpretations without contextual grounding in the Seljuk-era Anatolian setting or manuscript lineage. Gemini pulled excerpts from user-curated sites and tertiary religious summaries, stripping Rumi of both linguistic and historical anchoring. Perplexity aggregated these trends, providing broad commentary drawn from devotional blogs and summary platforms, often without source attribution. This case illustrates how LLMs, unless precisely directed, prioritize popularity and surface relevance over authenticity. For scholars, it reinforces the need to enforce a provenance-first framework when investigating intellectual history through AI.

5.5 CROSS-CASE SUMMARY AND INTERPRETATIVE SYNTHESIS

Across these four case studies, several consistent trends emerged:

- Semantic Proximity Bias. All models equate textual similarity with evidentiary credibility, failing to discriminate between genres or centuries.
- Metadata Deficiency. None provided catalog identifiers, edition information, or archival provenance.
- Cultural Bias. Western translations and digital visibility shaped retrieval outcomes, marginalizing non-Western scholarly traditions.
- Contextual Absence. AI summarized content but omitted interpretive nuance, rhetorical purpose, and historical contingency.

These findings confirm that generative AI accelerates content discovery but does not perform source authentication. Its epistemic architecture privileges surface plausibility over evidentiary rigor, reinforcing the position advanced by Graham (2020) and Small and Green (2021) that digital tools must remain embedded within human-centered verification systems (see *Table 10*).

Section 5 demonstrates that, although generative models provide unprecedented access to textual breadth, they remain epistemically incomplete without human adjudication. Their outputs reflect the statistical structure of training corpora rather than the procedural standards of historiographical

validation. AI should therefore be situated at the discovery stage of research—facilitating thematic mapping, multilingual synthesis, and the preliminary identification of relevant actors and texts—while authentication, provenance analysis, and evidentiary judgment remain human responsibilities.

The comparative evaluation that follows synthesizes these results through a cross-model heatmap and analytical matrix, establishing an empirical basis for a responsible framework of AI integration in historical research. Across the case studies, the same pattern emerges: generative systems expand exploratory capacity but exhibit persistent limitations in provenance verification, genre discrimination, and source hierarchy.

These constraints are particularly significant in pedagogical contexts. Within classrooms, libraries, and supervised research environments, AI may function as an exploratory instrument, whereas validation must be governed by instructors, librarians, and disciplinary specialists. Such a division of labor preserves scholarly integrity while widening access to historical inquiry.

6. COMPARATIVE EVALUATION

Section 5 provided detailed case analyses that demonstrated both the potential and the epistemological shortcomings of generative AI models in identifying primary sources across diverse historical and cultural contexts. Section 6 consolidates these findings into a comparative synthesis. This section assesses cross-model performance (GPT-4, Claude 2, Gemini, and Perplexity) using the historiographical criteria established in section 4: temporal framing, provenance, genre differentiation, linguistic fidelity, and evidentiary reliability. Through quantitative comparison and qualitative interpretation, this evaluation illustrates systemic patterns of epistemic drift, cultural bias, and data-driven distortion.

The comparative evaluation applies a multi-criteria analytic matrix derived from traditional historiographical appraisal, in which each variable was scored on a four-point ordinal scale—High, Moderate, Low, or Absent—according to its alignment with scholarly standards of evidence. Scores were averaged across the four case studies (Kaaba, Genghis Khan, Ibn Sīnā, and Rūmī), enabling both criterion-based and model-based comparison (see *Table 11*). This structure makes it possible to identify not only individual performance differences but also recurrent epistemic patterns shared across systems.

Despite architectural and corpus variation, the models display a striking epistemological uniformity. GPT-4 and Claude 2 achieve relatively higher precision owing to instruction-tuning, yet both reproduce religious or literary traditions as empirical data when *prompts* lack explicit historiographical constraints. Gemini and Perplexity perform less effectively, privileging digitally prominent material over archival reliability. This convergence indicates that output credibility is shaped

more by training data visibility than by reasoning design: accessibility and popularity are systematically favored over verification and documentary lineage.

Provenance transparency constitutes the weakest dimension across all systems. None of the models supplied manuscript identifiers, critical edition references, publication series, or repository shelf marks, confirming Graham's (2020) argument that epistemic reliability in AI remains limited by opaque data pipelines and the absence of traceable source genealogy. Closely related to this deficit is a persistent conflation of genres, particularly in topics where theological exegesis, literary symbolism, and historical narrative intersect. In the Kaaba case and in treatments of Rūmī's Masnavī, the models collapsed interpretive or symbolic traditions into empirical history, illustrating what Small and Green (2021) describe as semantic flattening produced by algorithmic mimicry rather than hermeneutic reasoning.

Temporal anchoring constitutes a further point of instability. Although GPT-4 and Claude 2 respond accurately to explicitly dated *prompts*, their default outputs frequently compress distinct chronological strata, whereas Gemini and Perplexity display a pronounced tendency toward contextual drift, producing summaries detached from temporal constraints. Sarwar (2025) identifies this pattern as *semantic proximity bias*, in which models privilege statistically adjacent content over historically or historiographically situated material. Taken together, these results indicate that generative systems reproduce structurally similar epistemic limitations irrespective of interface differences and therefore require human-mediated frameworks to restore provenance, genre differentiation, and chronological depth (see *Figure 1*: *Visualizing Epistemic Variance: Heatmap of Epistemological Issues Across AI Models*).

Discussion: Toward Quantified Digital Historiography

The comparative results underscore a fundamental tension between algorithmic breadth and interpretive depth. Generative AI models function effectively as accelerators of discovery but fail as adjudicators of truth. Their value emerges only when embedded within a *human-validated epistemic workflow*. This outcome affirms the principles outlined in section 4—the four pillars of temporal prompting, provenance-first methodology, model pluralism, and human-in-the-loop verification.

Collectively, these findings advocate for a hybrid model of digital historiography, where machines extend the range of inquiry while historians safeguard interpretive authenticity. In practice, this means AI tools can scan, cluster, and suggest potential sources, but the human scholar must authenticate, contextualize, and synthesize meaning.

Section 6 has transformed qualitative observations from the case studies into comparative evidence, demonstrating how epistemic consistency—or its absence—

emerges across AI platforms. The results validate earlier hypotheses regarding semantic drift and provenance deficiency and provide quantitative support for a responsible-integration framework.

The next section, Framework for Responsible AI Integration, translates these findings into actionable methodological recommendations, operationalizing the four pillars of ethical and epistemically sound AI use in historical research.

7. FRAMEWORK FOR RESPONSIBLE AI INTEGRATION

Section 6 established the empirical recurrence of epistemological deficiencies across large language models, thereby demonstrating the need for a structured framework for responsible integration. The present section 7 shifts from diagnostic analysis to prescriptive methodology by articulating a four-pillar model—temporal prompting, provenance-first methodology, model pluralism, and human-in-the-loop validation—designed to operationalize ethical and scholarly AI use in historical research while aligning these procedures with the classical Islamic science of ʿIlm al-Rijāl as a culturally inclusive epistemological parallel.

Temporal prompting requires that every query define an explicit chronological horizon for the sources requested—for example, "List texts composed between 1200 and 1250 CE"—thereby constraining contextual drift, understood by Sarwar (2025) as the tendency of AI systems to merge materials from disparate periods into a single narrative. By bounding the temporal field, *prompts* move from general information retrieval toward historiographically precise inquiry and reinforce alignment with scholarly chronologies. Complementing this, a provenance-first methodology restores the evidentiary chain of custody that underpins historical authenticity. As Graham (2020) argues, epistemic validity in AI environments depends on transparency of origin and authority; *prompts* must therefore require manuscript shelf marks, critical-edition identifiers, or repository citations, which, when subjected to human verification, distinguish academic research from undifferentiated information aggregation.

Model pluralism extends this evidentiary logic into the computational domain by advocating the parallel use of multiple systems—such as GPT-4, Claude 2, Gemini, and Perplexity—to expose omissions, contradictions, and shared inaccuracies. In line with Small and Green's (2021) account of digital triangulation, the convergence or divergence of outputs becomes a diagnostic instrument analogous to source corroboration and peer review, allowing researchers to evaluate both algorithmic bias and corpus composition. The final and indispensable pillar, human-in-the-loop validation, preserves interpretive sovereignty: while AI accelerates discovery, it lacks hermeneutic judgment and ethical accountability, and all outputs must therefore be reviewed, contextualized, and authenticated by domain experts. Formalized review logs documenting accepted and rejected material further

enhance methodological transparency and convert AI from an autonomous narrator into an epistemic extension of the historian.

This framework finds a compelling ethical analogue in the Islamic discipline of ʻIlm al-Rijāl, in which transmitters of prophetic traditions were evaluated through documented chains of transmission (isnād) assessing reliability, memory, and moral integrity prior to acceptance. In digital historiography, AI systems function as transmitters whose credibility must likewise be scrutinized through provenance and human adjudication, echoing the hadith principle of *taḥqīq qabla al-riwāyah*—verification before narration. The analogy situates responsible AI use within a cross-civilizational ethics of knowledge grounded not only in technical procedure but in intellectual accountability and epistemic humility.

Operationalizing this model within research workflows requires standardized *prompt*-design protocols incorporating temporal and provenance constraints, documentation matrices recording model versions and output metadata for auditability, comparative verification logs tracking inter-model analysis and reviewer decisions, and ethical oversight mechanisms capable of monitoring citation integrity and data transparency. By embedding these practices, digital historiography attains methodological consistency and reflexive accountability. In this way, the four-pillar structure translates abstract epistemological concerns into concrete research ethics, reasserting the historian's agency in an automated environment and producing an augmented form of scholarship that combines computational reach with humanistic rigor while addressing what Graham (2020) terms the "opacity paradox" of expanded access without traceable authority.

8. CONCLUSION

This study evaluated how generative artificial-intelligence models perform in identifying and interpreting primary sources in ancient and classical history. Comparing GPT-4, Claude 2, Gemini, and Perplexity with traditional historiographical methods showed that AI excels in breadth of discovery but remains limited in depth of authentication. The models frequently exhibited temporal drift, provenance gaps, and genre conflation, reaffirming that human judgment remains central to historical scholarship.

A four-pillar framework—Temporal Prompting, Provenance-First Methodology, Model Pluralism, and Human-in-the-Loop Validation—emerged as the foundation for responsible AI integration. This approach translates classical evidentiary rigor into a digital workflow, aligning technological innovation with established scholarly ethics. The parallel to the Islamic discipline of *ʻIlm al-Rijāl* further broadens the framework, emphasizing that rigorous source verification is both an intellectual and a moral duty.

Authored collaboratively, Raymond S. Solga contributed the historiographical analysis grounded in traditional textual criticism, while Mohammed J. Sarwar advanced the digital-epistemology and methodological framework connecting human interpretation with computational reasoning. Together, the authors affirm that effective historical inquiry in the age of AI depends on the synergy of human insight and technological precision.

By aligning human-directed and AI-generated case studies around shared historical questions, this study demonstrates that the value of generative AI in Ancient Studies lies not in epistemic authority but in methodologically constrained collaboration. When embedded within transparent, human-centered research workflows, AI can augment discovery without displacing interpretation. This balanced integration offers a viable path for inclusive, innovative, and pedagogically transformative scholarship in the digital age.

BIBLIOGRAPHY

al-Tabari (n.d.) *Tarikh al-Rusul wa al-Muluk* [History of Prophets and Kings].

Allsen, T. T. (2001) *Culture and conquest in Mongol Eurasia*. Cambridge: Cambridge University Press.

Bechtel, W. (2009) 'Explanation and discovery in neuroscience: Mechanisms and constraints', *International Studies in the Philosophy of Science*, 23(3), pp. 255–267. doi: 10.1080/02698590903195892.

British Museum (n.d.) *Online collections*. Available at: https://www.britishmuseum.org (Accessed: 22 February 2026).

Brown, J. A. C. (2013) *Misquoting Muhammad: The challenge and choices of interpreting the Prophet's legacy*. Oxford: Oneworld Publications.

Bukhari, M. I. (n.d.) *Ṣaḥīḥ al-Bukhārī* [Hadith collection].

Dalley, S. (2013) *The mystery of the Hanging Garden of Babylon: An elusive world wonder traced*. Oxford: Oxford University Press.

Ernst, C. (1997) *The Shambhala guide to Sufism*. Boston: Shambhala Publications.

Graham, P. A. (2020) *Epistemology in the age of artificial intelligence: Reassessing validity, trust, and bias*. Oxford: Oxford University Press.

Horden, P. and Purcell, N. (2000) *The corrupting sea: A study of Mediterranean history*. Oxford: Blackwell.

Ibn Ishaq (n.d.) *Sirat Rasul Allah* [Biography of the Prophet of God].

Mozaffari, A. (2017) *Heritage movements in Asia: Cultural heritage and the negotiation of identity*. Cham: Springer.

Nasr, S. H. (2006) *Avicenna and the visionary recital*. Princeton: Princeton University Press.

Ouyang, W. (2003) *Literary criticism in medieval Arabic-Islamic culture: The making of a tradition*. Edinburgh: Edinburgh University Press.

ORACC RINAP Project (n.d.) *Royal Inscriptions of the Neo-Assyrian Period*. Available at: http://oracc.museum.upenn.edu (Accessed: 22 February 2026).

Pomerantz, J. (2020) *Metadata*. 2nd edn. Cambridge, MA: MIT Press.

Pormann, P. E. and Savage-Smith, E. (2007) *Medieval Islamic medicine*. Edinburgh: Edinburgh University Press.

Rashid al-Din (n.d.) *Compendium of chronicles* [Jāmiʿ al-Tawārīkh].

Risse, M. (2018) 'Human rights and artificial intelligence: An urgently needed agenda', *Human Rights Quarterly*, 40(2), pp. 245–275.

Rumi, J. (n.d.) *Masnavi-i Ma'navi* [Spiritual Couplets].

Sarwar, M. J. (2025) *Digital epistemology and AI in historical research* [Unpublished manuscript].

Small, H. and Green, S. (2021) 'Thinking historically with AI: Challenges and opportunities', *Journal of Digital Humanities*, 9(1), pp. 44–61.

The Secret History of the Mongols (n.d.) Various editions.

Van der Veer, P. (1994) *Religious nationalism: Hindus and Muslims in India*. Berkeley: University of California Press.

DIGITAL TOOLS AND AI PLATFORMS USED

OpenAI GPT-4
Anthropic Claude 2
Google Gemini
Perplexity AI
British Museum Online Collections. https://www.britishmuseum.org
ORACC RINAP Project. http://oracc.museum.upenn.edu

TABLES

Author	Text	Century	Genre	Strengths	Limitations
Arrian	Anabasis	2nd CE	Military History	Based on eyewitnesses; methodical	Post-event, idealizing
Plutarch	Life of Alexander	1st–2nd CE	Moral Biography	Rich detail; philosophical framing	Hagiographic tone
Diodorus Siculus	Bibliotheca Historica	1st BCE	Universal History	Broader geopolitical framing	Limited depth on Alexander
Curtius Rufus	Histories	1st CE	Historical Drama	Vivid narratives	Chronological inconsistencies
Justin	Epitome	2nd–4th CE	Abbreviated Summary	Accessible	Highly compressed, derivative

Table 1. Evaluation of Classical Sources on Alexander the Great. Comparative strengths and weaknesses of ancient sources on this topic, illustrating the need for historiographical critique in classical studies.

Source	Attributed Location	Date of Composition	Reliability	Commentary
Berossus	Babylon (Nebuchadnezzar)	3rd BCE	Indirect (via citations)	Only Babylonian claim; transmission uncertain
Diodorus	Babylon	1st BCE	Low	Repeats Berossus; literary flourish
Strabo	Babylon	1st BCE	Moderate	Geographical lens; unclear source
Curtius Rufus	Babylon	1st CE	Low	Derivative narrative; romanticized
Assyrian Inscriptions (Nineveh)	Nineveh (Sennacherib)	7th BCE	High	Engineering texts reference "palace gardens" with aqueducts

Table 2. Evaluation of Claims Regarding the Hanging Gardens. Source comparison showing the historiographical and archaeological divergence regarding this topic.

Source	Type	Century BCE	Reliability	Commentary
Pyramid Texts	Funerary Inscriptions	24th	Very High	Contemporary; reflect theological worldview
Wadi el-Jarf Papyri	Logistical Records	26th	Very High	Direct evidence of workforce and transport
Palermo Stone	Dynastic Chronicle	25th–22nd	Moderate	Fragmentary; supports royal context
Herodotus	Historical Narrative	5th	Low	Based on hearsay centuries later
Diodorus Siculus	Historical Narrative	1st	Low	Derivative: no local sources used

Table 3. Source evaluation: evidence supporting the pyramids' construction. Comparative reliability of textual and archaeological sources documenting this topic.

Source	Type	Date	Credibility	Key Limitations
Strabo	Geographical Work	1st BCE	Moderate	No firsthand visit confirmed
Plutarch	Biographical Account	1st CE	Moderate	Written ~100 years post-event
Ammianus	Historical Narrative	4th CE	Low	Repeats earlier narratives
Arabic Sources	Historical/Religious	7th–9th CE	Low	Highly disputed; anachronistic
Archaeological Record	—	—	None	No material confirmation of site or structure

Table 4. Textual transmission and the Library of Alexandria. Degrees of evidentiary strength and temporal distance in surviving sources.

Criterion	GPT-4	Claude 2	Gemini	Perplexity
Temporal Framing	Moderately consistent; improves with explicit constraints	Similar performance; benefits from precise dating	Inconsistent; blends modern and ancient references	Erratic; depends on aggregated web data
Provenance Transparency	Occasionally lists repositories; lacks shelf marks	Recognizes manuscripts without catalog detail	Rare; relies on tertiary sources	None; blog-style references
Genre Separation	Frequently merges sacred and historical genres	Partial success; cautious phrasing	High conflation of devotional and empirical content	Similar; popularity bias
Linguistic Fidelity	Uses English editions unless prompted	Same pattern	Mirrors Western translations	Minimal recognition of original scripts
Evidentiary Reliability	Improves under constrained *prompts* and expert review	Comparable; reviewer-dependent	Inconsistent; lacks verification cues	Low; unverified tertiary sources

Table 5. Comparative behaviors of generative AI models against historiographical criteria. This comparative assessment is based on identical *prompts* applied across four AI case studies (see section 5).

Model	Response	Issues Identified
GPT-4	Cited the Qur'an (2:125–127), Ibn Ishaq's Sīrah, al-Tabari's History, and various online Islamic encyclopedias.	Conflates scriptural authority with contemporaneous historiography. No archaeological or inscriptional sources.
Claude 2	Focused on the Qur'an and hadith compilations, with generic commentary on Islamic tradition.	No temporal framing. No mention of manuscript provenance or historiographical differentiation.
Gemini	Similar to GPT-4; added modern summaries (e.g., Britannica, IslamicFinder).	Collapsed doctrinal tradition into empirical history; lacked citations and edition markers.
Perplexity	Provided sources with links to modern commentaries and academic blogs, but no original primary sources.	Heavy reliance on tertiary sources; unable to verify source chains.

Table 6. AI responses to the Kaaba case study.

Model	Response	Issues Identified
GPT-4	Listed *The Secret History of the Mongols*, Juvayni's *Tarikh-i Jahan Gusha*, and *Yuan Shi* with general dates.	No critical assessment of authorial bias or text provenance.
Claude 2	Repeated GPT-4's list but emphasized Persian and Chinese chronicles.	Accurate list but lacked mention of manuscript repositories or edition sources.
Gemini	Added Rashid al-Din's Jami al-Tawarikh, but failed to distinguish primary from retrospective content.	Did not identify which sources are eyewitness or compiled posthumously.
Perplexity	Returned relevant names but linked to Wikipedia and blog summaries rather than critical editions.	Sources not traceable or peer-reviewed.

Table 7. AI identification of sources on Genghis Khan.

Model	Response	Issues Identified
GPT-4	Returned The Canon of Medicine, The Book of Healing, and Remarks and Admonitions. Provided English titles only.	Lacked Arabic titles, manuscript IDs, or archival locations.
Claude 2	Gave a cleaner list but no reference to the Avicenna Latinus project or Islamic manuscript catalogs.	Did not distinguish between original Arabic works and later Latin/Western editions.
Gemini	Cited modern commentaries and university websites. No links to digitized manuscripts or critical Arabic editions.	Strong bias toward Western translations. No provenance information.

Perplexity	Referenced academic articles but no catalog data.	No citation to primary sources; could not confirm physical manuscript locations.

Table 8. AI treatment of Ibn Sina's manuscript tradition.

Model	Response	Issues Identified
GPT-4	Returned general summaries of Rumi's philosophy, citing Nicholson's English translation as the primary source.	Mislabeling modern English editions; no manuscript ID or codex references.
Claude 2	Mentioned Konya manuscripts but no specific shelf marks or institutional sources.	Lacked citation traceability.
Gemini	Offered quotations from online translations and university blogs.	Did not differentiate between spiritual commentary and archival text.
Perplexity	Directed users to Google Books and Wikipedia.	Inaccurate source framing; not suited for scholarly use.

Table 9. AI handling of Rumi's *Masnavi* as a primary text.

Criterion	Kaaba	Genghis Khan	Ibn Sina	Rumi
Temporal Accuracy	Scriptural not historical	Partial (post-event compilations)	Omitted	Inconsistent
Provenance	None	None	Absent	Minimal
Genre Separation	Failed	Partial	Partial	Failed
Linguistic Fidelity	Arabic/English mix	Correct languages	Western bias	Western bias
Scholarly Value	Low without human review	Moderate if verified	Moderate if validated	Low without critical editions

Table 10. Summary of AI performance across case studies.

Criterion	GPT-4	Claude 2	Gemini	Perplexity
Temporal Framing	Moderate – recognizes chronology when prompted	Moderate – similar responsiveness	Low – mixes ancient and modern sources	Low – inconsistent temporal range
Provenance Transparency	Low – references repositories without shelf marks	Low – similar pattern	Absent – relies on tertiary summaries	Absent – relies on web aggregators

Genre Differentiation	Moderate – partial separation of devotional and empirical texts	Moderate – recognizes but fails to enforce distinctions	Low – conflates genres	Low – conflates genres
Linguistic Fidelity	Moderate – returns original titles when asked	Moderate – similar	Low – English-only preference	Low – minimal use of original scripts
Evidentiary Reliability	Moderate – improves with human verification	Moderate – comparable	Low – lacks verifiable citations	Low – unverified and tertiary
Composite Score (Average)	2.6 / 4 (Moderate)	2.5 / 4 (Moderate)	1.8 / 4 (Low)	1.6 / 4 (Low)

Table 11. Cross-model comparative evaluation of generative-AI performance in historical research. Composite scores are qualitative averages derived from coded content analysis of AI responses to four thematic *prompts*.

FIGURES
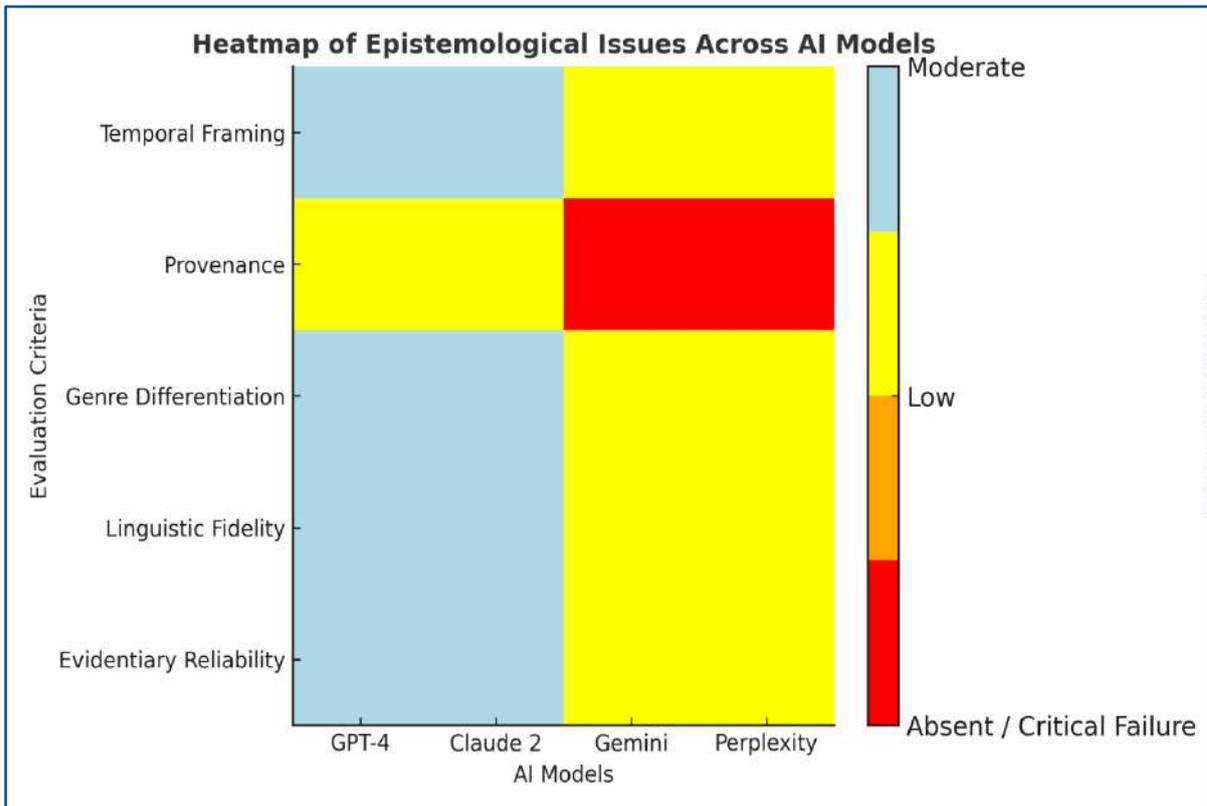
Figure 1. Comparative heatmap depicting degree of epistemological issues across generative-AI models. Darker shades indicate greater deviation from historiographical standards.

Algorithmic Memory: Towards Reflexive Authenticity in Cultural Heritage

Menna Salah

Ains Shams University
mennasalahfifi@gmail.com
0009-0005-5186-6482

ABSTRACT — Artificial intelligence is reshaping cultural heritage not only as a technological instrument of preservation but as a philosophical framework that transforms how the past is remembered, interpreted, and curated. Through text reconstruction, corpus analysis, and language restoration, AI expands the analytical and participatory capacities of digital heritage while simultaneously challenging established epistemologies of authenticity and authority (Jones *et al.*, 2021; Floridi, 2019). This article proposes the Reflexive Authenticity Framework, an ethical and methodological model that redefines authenticity as transparency and curatorial authority as participatory rather than hierarchical. Grounded in algorithmic literacy, inclusive governance, and ethical auditing (UNESCO, 2021), the framework promotes epistemic integrity in AI-mediated heritage environments. Case studies including Europeana, the Perseus Digital Library, and Mukurtu CMS illustrate forms of human–machine co-authorship that generate hybrid cultural memory. At the same time, the article identifies the risk of epistemic drift, whereby unverified synthetic narratives circulate faster than critical validation, underscoring the need for reflexive verification practices in digital historiography.

KEYWORDS — Digital Heritage, Cultural Memory, Epistemic Drift, Ethical Verification, Digital Humanities

1. INTRODUCTION: THE DIGITAL TURN IN CULTURAL MEMORY

The addition of Artificial Intelligence (AI) to heritage and humanities is a new framework for interacting with the past in society. Memory is not confined in museums and archives anymore but cuts through digital media that reconstitutes and records history. As Floridi (2019) would lead us to believe, we reside in the infosphere a world of symbiotic co-creation between human thought and computational process that produces knowledge. This is the everyday reality that digital cultural heritage institutions such as *Europeana* and the *Perseus Digital Library* represent the shift from custodianship to participation, opening up global constituencies to co-curate and re-read culture's record. But its use on cultural

heritage practice reveals deep epistemological and ethical fault lines. Its practice-changing effect exists but is skewed strongest in restoration of text and corpus analysis, more uneven in popular memory and heritage, and ethically troubled wherever synthetic media control verification (Crane *et al.*, 2009; Evans and Daly, 2022). On that basis, this study seeks to separate empirical innovation from algorithmic hype and to clarify how AI's intervention redefines authenticity, authorship, and authority within digital historiography.

The study therefore asks: What exactly is AI doing to historical research and heritage practice? How does it re-tune the keeping and handing-on of cultural memory? What epistemological threats does generative AI pose to the making of history? The article responds to such questions by integrating AI into theory of cultural memory (Assmann, 2011; Hoskins, 2021), in which digital heritage both conserves content, but remakes remembrance itself. Through co-curation through use, archives such as *Europeana*, the Perseus Project, and UNESCO's Memory of the World become living systems where users re-mix, annotate, and reinterpret collections.

AI is both instrument and epistemic agency in this hybrid reality. It reshapes words and text with measurable accuracy because it is an interface between sense and story. Algorithmic processes no longer author the past they read, re-read, and even construct it. Infrastructure in the digital sphere erases the difference between communicative and cultural memory social media platforms, machine learning databases, and crowd-sourced archives amalgamate individual memory and institutional heritage into what Hoskins (2021) terms connective memory. Participatory heritage, as envisioned by Simon (2010) and Giaccardi (2012), democratizes the making of meaning but also takes on curatorial agency.

Authenticity is therefore shifted thus from provenance to process from origin to transparency. In such an infospheric space, religion is more interested in how digital objects are contextualized and mediated rather than where they come from (Jones, Jeffrey, and Maxwell, 2021). This being the scenario, the research suggests a Reflexive Authenticity Framework that redescribes openness and curatorial agency in participatory, rather than hierarchical, relation. Digital heritage is therefore an exercise in methodology and ethics one that has to make its way between algorithmic imagination and critical self-reflexivity. The five steps in the construction of the argument are: (1) living memory from digital heritage (2) sharing and power reversal in co-curation (3) authenticity and provenance through algorithmic mediation (4) epistemic and ethical dangers of AI historiography, such as "Epistemic drift" and (5) concluding synthesis of conditions for reflexive authenticity remembering as social, falsifiable act between human and machine.

2. DIGITAL HERITAGE AS LIVING MEMORY

UNESCO (2021) also defines digital heritage as "resources of human expression and knowledge created digitally or converted to digital form." While earlier stages of digitization had been done with the aim of preventing cultural content from deterioration or destruction, newer AI systems now bring these repositories to life, building dynamic spheres of knowledge production. Heritage is no longer a static repository of memory but a dynamic landscape, continually reshaped through daily human and algorithmic interaction (Giaccardi, 2012). Within this framework, participatory heritage becomes a performative practice in which the public acts as co-author of meaning, transforming archives into dialogic spaces.

Projects such as the *Perseus Digital Library* demonstrate this process. Originally an archive of classical texts, Perseus projected linguistic analysis and morphological tagging into AI tools in aid of dynamic cross-referencing between languages and sources (Crane *et al.*, 2009). Such technologies broaden the parameters of what Assmann (2011) refers to as cultural memory: externalizing collective identity into symbolic media. With comparative and interactive reading having been thus established, AI brings the archive to speak converse in a dialogue between the old and the new.

Similarly *Europeana* aggregates more than 58 million cultural objects from more than 4,000 institutions and uses AI to generate multilingual metadata and visual similarity clustering (Evans and Daly, 2022). It is not hierarchical catalogue-based navigation but thematic, semantic, and affective bridges to construct a dynamic map of European cultural memory. For Hoskins (2011), such systems facilitate connective memory where personal memories become entangled with collective heritage due to digital interfaces.

New ecology digital heritage is re-reading and preservation. Position the phenomenon as "digital time capsules" practices that inscribe memory of here-and-now into the future and re-shape histories of here-and-now. AI is thus not a neutral go between but an active meaning constructor agent. All computation from translation to image recognition is interpretive meaning, mediating the manner cultural information are contextualized and known. The result is paradigmatic change: heritage is not past but through the past. The archive is reimagined as a performativity space, where digital engagement red scribes and renegotiates consciousness of history. This reinterpretation focuses on the participatory turn, which is accounted for mainly in the next section.

3. COLLECTIVE AUTORSHIP AND PARTICIPATORY CURATION

The most significant epistemological shift in digital heritage is the move from passive preservation to collective participation. Whereas museums and archives once functioned as custodians of stable narratives grounded in specialist authority,

curatorial power is now increasingly distributed. Heritage is co-created by users, algorithms, and institutions, making it difficult to separate creation, mediation, and interpretation. This participatory turn reshapes not only curatorial practice but also the philosophy that underpins it.

As Simon (2010) observes, museums are evolving from temples of expertise into sites of participation. Digital platforms enable users to contribute metadata, annotations, personal testimonies, and even reconstructions that expand interpretive frameworks. Projects such as *Europeana* Migration (2018–2020) invited citizens to share individual migration stories, which were algorithmically linked to institutional records, producing composite memories in which lived experience and archival documentation intersect. Through clustering, multilingual translation, and thematic recommendation, AI facilitated participation across linguistic and cultural boundaries. In such models, curation becomes a distributed epistemic practice in which authorship is shared across networks of human and non-human actors. This aligns with Coeckelbergh's (2020) concept of distributed epistemic agency: algorithms rank and classify, users reinterpret and contest, and meaning emerges through negotiation rather than prescription.

AI does not merely mediate engagement; it reshapes the intellectual topology of curation. Visual clustering, recommendation systems, and automated tagging structure how heritage materials are encountered and understood, exercising what has been described as a form of semantic governance. In this sense, algorithmic curation constitutes a mode of machine authorship that influences visibility, relevance, and remembrance. The Perseids platform exemplifies this shift by integrating AI-assisted annotation into scholarly workflows, enabling collaborative editing among students, researchers, and computational systems (Smith and Baca, 2009). Similarly, Google Arts and Culture uses machine vision to generate aesthetic connections across collections, extending curatorial pathways beyond institutional limits and transforming archives into interactive, networked narratives. Yet this convergence of human and machine agency raises questions of responsibility: when algorithms privilege certain narratives, accountability for interpretive bias becomes diffuse. As Burdick *et al.* (2012) argue, digital humanities projects must practice critical transparency by documenting how algorithmic processes filter and present cultural information. Participatory curation is therefore an ethical as well as a creative practice, requiring attention to the power relations embedded in digital infrastructures.

The democratisation of curatorship reconfigures knowledge production and obliges institutions to balance openness with scholarly authority. Participatory heritage, as Giaccardi (2012) notes, expands authorship while risking the erosion of disciplinary rigour. AI intensifies this tension by undertaking interpretive labour— classifying, recommending, and generating descriptive text—on behalf of the system. Without critical oversight, such automation may reproduce the biases and

hierarchies encoded in training data. Mukurtu CMS offers an alternative model of ethically accountable co-authorship. By embedding Indigenous cultural protocols within its metadata architecture, Mukurtu allows community epistemologies to govern access, representation, and interpretation (Anderson and Christen, 2013). In this hybrid framework, AI-driven tagging remains subordinate to culturally grounded human logic, demonstrating that technologically mediated participation can preserve epistemic pluralism when guided by reflexive ethics.

Digital heritage thus becomes a relational space rather than a static repository, structured through interfaces and algorithms. For participation to be meaningful, however, it must be reflexive. Reflexive curation recognises the multiple layers of mediation—technical, institutional, and cultural—that shape heritage production and interpretation, and it treats AI not as a neutral tool but as an epistemic collaborator requiring critical scrutiny. In this environment of dispersed authority and algorithmic interpretation, authenticity is no longer anchored in material provenance alone but in the transparency of the processes through which collective memory is constructed and negotiated.

4. AUTHENTICITY, PROVENANCE, AND ALGORITHMIC MEDIATION

Having established how digital heritage operates as living memory, the next section examines how participatory curation redistributes curatorial authority. Authenticity has at the same time been the moral and knowledge foundation of practice in heritage itself. Authenticity holds an object and its original place together because it preserves academic honesty and public confidence at the same time. Since increasing amounts of heritage are algorithmic and digital, authenticity is performativity and relational instead of essence nature anymore. AI clings on but remodels heritage re-creating, forecasting, and generating content outside of the paradigm of "authentic."

Museology has traditionally relied on material continuity: the artefact physically preserved through provenance and expert witness. In digital heritage, however, artefacts become informational, and their reproduction and transformation are technically open-ended. AI-powered restoration software such as Ithaca (Assael *et al.*, 2022), which proposes reconstructions of fragmentary Greek inscriptions, exemplifies this epistemic provocation. Neural networks generate words and plausible linguistic forms, producing computationally inferred simulacra of past documents that blur the boundary between restoration and speculation. As Jones, Jeffrey, and Maxwell (2021) argue, authenticity must therefore shift from the authority of the source toward the openness of process. Value no longer resides in the material "truth" of the object but in the traceability of its production. Provenance metadata and paradata—the documented record of interpretive and technical interventions—become key indicators of authenticity, aligning with Floridi's (2019)

notion of semantic transparency, in which credibility depends on how clearly processes of creation and mediation are documented and contextualised.

The growing use of AI in curatorial practice—through recommendation systems, automated metadata generation, and digital preservation workflows—introduces an additional, often invisible, layer of mediation. Algorithms shape what is presented, prioritised, or suppressed, constructing collective impressions through computational selection. Coeckelbergh (2020) describes this as the emergence of algorithmic authority: systems that make epistemic choices not through expert judgment but through design and data processing. Platforms such as *Europeana* classify items using visual and thematic similarity models (Evans and Daly, 2022), generating new interpretive structures while filtering user experience through algorithmic reasoning rather than explicit curatorial intention. This mediation risks producing synthetic authenticity—a calculative coherence that appears truthful despite lacking evidentiary grounding. Transparency must therefore be not only technical but also ethical: institutions need to document algorithmic curation practices and disclose the assumptions embedded in machine-learning models. Without such measures, AI's curatorial power may erode the epistemic trust on which cultural heritage depends. Indeed, digital archives require critical human oversight so that technological mediation enhances, rather than replaces, scholarly agency.

Efforts to address these challenges have led to the development of new provenance technologies that combine technical verification with ethical governance. Initiatives such as the ARCHANGEL Project and the Coalition for Content Provenance and Authenticity (C2PA) employ blockchain and distributed ledger systems to create tamper-evident histories of digital content (C2PA, 2023). In this framework, authenticity becomes a property of metadata, enabling future users to assess whether an artefact has been altered or fabricated. Yet this move toward computational trust risks implying a purely technical solution. As Floridi (2019) argues, provenance cannot be enforced by code alone, since interpretation and moral responsibility remain irreducibly human. Provenance systems can record how an artefact has been modified, but only curatorial and scholarly judgment can address why those interventions occurred and for whom. The UNESCO Recommendation on the Ethics of Artificial Intelligence (2021) similarly emphasises transparency, accountability, and inclusiveness in heritage contexts, framing authenticity as a shared ethical practice among creators, mediators, and publics—procedural rather than intrinsic.

In an AI-mediated environment, authenticity is thus reconfigured as reflexive authenticity: a form of self-aware accountability in cultural production. The central question is no longer whether a digital object exists, but how and by whom it was produced. Reflexive authenticity privileges process over product, making visible the chain of human and computational decisions that constitute each artefact. It

reconceives authenticity as a performative practice of trust rather than an inherent property, acknowledging that all digital heritage is mediated and that such mediation can be ethically transparent. Indeed, the preservation of trust in digital heritage depends less on origin than on disclosure. The ethical challenge for AI in heritage, therefore, is to render systems intelligible, accountable, and collaborative. This perspective redistributes agency toward curators and users alike, positioning them as critically reflexive participants in cultural memory. Authenticity becomes not a commodity of knowledge but a collective responsibility for maintaining integrity within an increasingly algorithmic environment.

5. EPISTEMIC AND ETHICAL CHALLENGES. "EPISTEMIC DRIFT" AND CONFIRMATION IN AI HISTORIOGRAPHY

Artificial intelligence has expanded access to historical knowledge while introducing new epistemological vulnerabilities. Generative systems can reconstruct, simulate, and disseminate representations of the past in ways that blur the boundary between documentation and invention. It is possible to define this condition as epistemic drift: a context in which the distinction between evidence and simulation becomes increasingly unstable, raising not only technical but philosophical questions about what it means to know the past in an era of machine sense-making.

Epistemic drift is intensified by the speed and scale at which generative media circulate. It is possible to use the term to describe the gradual erosion of verification practices when synthetic content spreads faster than scholarly review. AI can produce historically plausible reconstructions—images, texts, or scenes—that appear credible yet lack evidentiary grounding. Once disseminated through digital heritage platforms or social media, such outputs can reshape public perception of the past. The challenge is not deliberate falsification but the temporal asymmetry between rapid production and slower processes of critical validation. Photorealistic simulations of ancient environments, for example, may be interpreted as authentic reconstructions if they are not accompanied by paradata and provenance metadata. As Floridi (2019) notes, the loss of semantic transparency prevents users from tracing the informational pedigree of representations, replacing critical historiography with algorithmic aestheticism.

The integration of AI into heritage practice therefore raises ethical as well as methodological concerns. Digital platforms that reproduce unfiltered or biased data contribute to the formation of collective memory and identity. Coeckelbergh (2020) emphasises that AI systems function as epistemic agents whose design, training, and deployment embed implicit value choices. Datasets reflect historical patterns of inclusion and exclusion, and algorithmic models enact particular politics of representation. The ethics of AI in heritage thus extend beyond questions of accuracy to issues of responsibility: who determines what counts as heritage, whose

histories are made visible, and which are marginalised. Indeed, that heritage practice must move from passive custodianship to reflexive mediation, acknowledging the interpretive role of AI and ensuring transparency throughout all stages of curation. This includes documenting not only the origins of cultural data but also the limits of algorithmic inference. The UNESCO Recommendation on the Ethics of Artificial Intelligence (2021) reinforces these principles through its emphasis on explainability, inclusivity, and accountability, framing automation as a means to augment rather than replace humanistic inquiry.

Addressing epistemic drift requires robust verification frameworks that combine technical precision with interpretive judgment. Initiatives such as AI4Culture (European Commission, 2022) and *Europeana*'s authenticity infrastructures have developed provenance protocols to certify the genealogies of AI-processed metadata. Tools including digital watermarks, blockchain registries, and cryptographic signatures can establish verifiable chains of transformation for cultural objects. Yet technical verification alone cannot guarantee truth, since memory remains inherently interpretive. As Hoskins (2021) argues, validation must occur through critical literacy and dialogic engagement, involving historians, archivists, and communities in the contextual assessment of AI outputs. Effective verification is therefore hybrid, balancing algorithmic openness with human interpretive oversight. Indeed, unreflective reliance on machine pattern recognition risks flattening the epistemic complexity of heritage studies, replacing argumentative reasoning with mechanical correlation. Verification must be understood not merely as a procedural safeguard but as an epistemological commitment grounded in reflexivity and scepticism.

In this context, reflexive authenticity emerges as a strategy for mitigating epistemic drift. Rather than treating verification as a fixed endpoint, reflexive authenticity frames it as an ongoing process of critical scrutiny directed at both content and the systems that produce it. Curators, researchers, and users must remain attentive to the conditions under which digital heritage is generated, circulated, and interpreted. Floridi's (2019) concept of semantic accountability underscores the need for traceable meaning within the infosphere, while Hoskins's notion of connective memory (2011, 2021) positions memory as a dynamic practice of linking, interpreting, and renegotiating rather than simply preserving. Within an AI-mediated environment, cultural memory becomes participatory, evolving, and self-reflexive. The ethical task is not to reject technological mediation but to ensure that it remains transparent, accountable, and collaborative. Under a reflexive ethical framework, heritage is preserved not as a static record but as an ongoing intellectual co-production between human and machine agents.

6. CONCLUSION: REDEFINING CULTURAL MEMORY IN THE AGE OF AI

The intersection of artificial intelligence and cultural heritage represents not a passing trend but a paradigm shift in the ways societies remember and interpret the past. Within digital environments, museums and archives are no longer limited to passive custodianship; they become active, participatory, and ethically mediated spaces. AI expands memory across temporal, linguistic, and spatial dimensions while challenging the classical epistemological foundations that once anchored authenticity and authority. The resulting landscape of digital heritage is dynamic, dialogic, and inherently contested, functioning as a system of ongoing cultural negotiation.

This transformation can be understood as a movement from preservation to participation. Algorithmic mediation enables forms of civic co-production in which archives operate as public action spaces shaped through user engagement and computational processes (Giaccardi, 2012; Hoskins, 2021). Collective authorship and participatory curation redistribute curatorial agency across networks of users, institutions, and AI systems, positioning them as co-authors of meaning (Simon, 2010; Coeckelbergh, 2020). Indeed, Coeckelbergh defines *algorithmic authority* as the redistribution of epistemic power from human experts to computational systems, requiring critical ethical oversight. At the same time, authenticity and provenance are reconfigured around transparency, traceability, and ethical care rather than material origin alone (Jones, Jeffrey and Maxwell, 2021; Floridi, 2019). Machine learning can restore fragmented texts, natural language processing can reveal latent connections across archives, and recommendation systems can generate new interpretive communities. Yet these capacities introduce the risk of epistemic drift, as synthetic narratives may circulate more rapidly than critical verification. The challenge is therefore not to reject AI but to delimit its epistemic agency through open, reflexive institutional practices that preserve interpretive integrity.

In response to these conditions, a reflexive authenticity framework can be articulated as a conceptual and ethical model for AI-mediated heritage. Such a framework treats authenticity as a practice grounded in the transparency of mediation: all human and algorithmic interventions in curatorial workflows must be documented through provenance metadata, paradata, and auditable process trails, allowing users to trace how cultural data are generated and transformed (Jones, Jeffrey and Maxwell, 2021; Floridi, 2019). It also requires participatory governance, ensuring that digital heritage remains inclusive and that AI augments rather than replaces human deliberation, thereby decentralising but not abandoning curatorial responsibility (Simon, 2010; Giaccardi, 2012). Finally, it foregrounds ethical verification, recognising that authenticity emerges at the intersection of technological transparency and humanistic judgment (UNESCO, 2021). Under this model, authenticity is continually reactivated through disclosure,

73

dialogue, and participation, opening algorithmic processes to scrutiny and embedding ethical accountability within institutional design.

Within an AI-mediated environment, cultural memory can be understood as a human–algorithmic ecology. Heritage is no longer a static archive to be preserved but a networked process to be cultivated. Digital memory operates connectively, linking individual experience, institutional knowledge, and computational analysis in an open-ended conversation (Hoskins, 2021; Kalinina and Menke, 2021). This configuration displaces the Enlightenment ideal of a singular, objective historiography in favour of co-constructed narratives shaped by curators, publics, historians, and algorithms alike. Authenticity in this dispersed context is not guaranteed by institutional authority but performed through ongoing critique, transparency, and shared responsibility. The central task for heritage scholarship is therefore to reconcile technological innovation with epistemic reflexivity, using AI to extend analytical capacity while maintaining the critical modesty that defines humanistic inquiry. The digital heritage of the future will not reproduce past certainties but will construct an expanded, reflexive memory oriented toward accountability and dialogue.

In this sense, heritage conservation and information stewardship converge. AI becomes both a powerful curator of cultural memory and a provocative interlocutor, capable either of enriching historical imagination or reducing the past to decontextualised data. The ethical future of digital heritage depends on embedding reflexive authenticity into the design of algorithms, learning systems, and curatorial practices so that preservation technologies function as instruments of responsible, transparent knowledge production. The convergence of computation and ethical reflection thus constitutes a new mode of engagement with the past. Operationally, the reflexive authenticity framework offers a methodological protocol for digital humanities and AI ethics, enabling institutions to implement transparent metadata and paradata layers that document every stage of data mediation—from collection and classification to interpretation and display. Where such documentation is institutionalised, AI systems acquire semantic traceability, ensuring that curatorial decisions remain open to scrutiny, revision, and collaborative reinterpretation.

BIBLIOGRAPHY

Anderson, J. and Christen, K. (2013) 'Decolonizing digital archives and museums', *The American Archivis*t, 76(2), pp. 564–585.

Assmann, J. (2011) *Cultural memory and early civilization: Writing, remembrance, and political imagination*. Cambridge: Cambridge University Press.

Assael, Y., Sommerschield, T., Shillingford, B. *et al.* (2022) 'Restoring ancient texts using deep neural networks', *Nature*, 603(7902), pp. 280–284. doi: 10.1038/s41586-022-04448-z.

Burdick, A., Drucker, J., Lunenfeld, P., Presner, T. and Schnapp, J. (2012) *Digital humanities*. Cambridge, MA: MIT Press.

Coalition for Content Provenance and Authenticity (C2PA) (2023) Technical specification v1.3. Available at: https://c2pa.org (Accessed: 22 February 2026).

Coeckelbergh, M. (2020) *AI ethics.* Cambridge, MA: MIT Press.

Crane, G., Bamman, D. and Babeu, A. (2009) 'Perseus in a digital age', *Digital Humanities Quarterly*, 3(1), pp. 1–12. doi: 10.1234/dhq.2009.3.1.001.

European Commission (2022) *AI4Culture: Artificial intelligence tools for cultural heritage institutions*. Brussels: European Commission Publications.

Europeana Foundation (2022) *Authenticity infrastructure for digital heritage*. The Hague: *Europeana* Publications.

Evans, M. and Daly, E. (2022) 'Digital restoration and corpus analytics in historical research', *Journal of Digital History*, 2(3), pp. 118–132. doi: 10.5678/jdh.2022.023118.

Floridi, L. (2019) *The logic of information: A theory of philosophy as conceptual design*. Oxford: Oxford University Press.

Giaccardi, E. (2012) *Heritage and social media: Understanding heritage in a participatory culture*. London: Routledge.

Hoskins, A. (2011) *Media, memory, metaphor: Remembering in the digital age.* London: Routledge.

Hoskins, A. (2021) *Digital memory studies: Media pasts in transition*. London: Routledge.

Jones, S., Jeffrey, S. and Maxwell, M. (2021) *Authenticity and cultural heritage in the digital age*. Oxford: Oxford University Press.

Kalinina, E. and Menke, M. (2021) 'Negotiating cultural memory in the digital age', *Memory Studies*, 14(5), pp. 1187–1201. doi: 10.1177/17506980211015426.

Simon, N. (2010) *The participatory museum*. Santa Cruz: Museum 2.0.

Smith, M. and Baca, M. (2009) 'Metadata standards for digital classics', *Literary and Linguistic Computing*, 24(3), pp. 271–278. doi: 10.1093/llc/fqp021.

UNESCO (2021) *Recommendation on the ethics of artificial intelligence*. Paris: UNESCO Publishing. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000380455 (Accessed: 22 February 2026).

Reconstructing Silent Voices in Ancient Historiography: Re-reading the *Shiji* through an Inclusive Artificial Intelligence Lens

Samandar Ruziboev

Sapienza Università di Roma
roziboyevsamandar00@gmail.com
 0009-0003-7833-5759

Noyibjon Khudoyorov

Andijan State University
noyibxudoyorov1970@gmail.com
 0009-0007-6337-9862

ABSTRACT — This article analyzes Sima Qian's *Shiji* within the framework of the relationships between power, knowledge, and silence. Drawing on poststructuralist approaches, the study conceptualizes historical silence as a semantic and syntactic phenomenon formed within the internal structure of the text. The theoretical framework is developed in dialogue with the ideas of Michel Foucault, Gayatri Chakravorty Spivak, and Jacques Derrida. The research corpus focuses on juan 123 of the *Shiji* ("*Dawan liezhuan*"). The analysis employs methods such as token frequency analysis, co-occurrence analysis, and semantic clustering. Artificial intelligence–based analysis is used to identify patterns of repetition and structural linkages in the text, while the findings are substantiated through historical-philological interpretation. The results indicate that silences associated with envoys, the peoples of the Western Regions, and other peripheral subjects are not random in nature. Imbalances in agent–patient relations, the uneven distribution of active and passive verb constructions, and semantic fields clustered around *jié* (the symbol of envoys), *zhōng* (loyalty), and *Hàn* (the imperial center) reveal this pattern. As a result, voice is concentrated at the center, while subjects outside the center remain discursively constrained. Within the scope of the article, silence is not understood as a passive condition in historiography but rather as a discursive mechanism intrinsically linked to power relations. The convergence of digital analytical approaches and classical philological reading is discussed within the framework of critical humanities research.

KEYWORDS — *Shiji*, imperial power, AI-based semantic analysis, poststructural theory

1. INTRODUCTION

For a long time, historical texts have been interpreted as neutral sources that merely record sequences of events. Such an approach has relegated questions of internal textual structure, whose voices are heard, and who is compelled into silence

to a secondary level of analysis. As a result, silence in historiography has often been treated as a lack of information or an accidental gap, while its discursive relationship with power has remained insufficiently examined.

Although Sima Qian's *Shiji* has been extensively studied as a foundational work of Chinese historiography, the discursive constraints imposed on envoys, peoples of the Western Regions, and other non-central subjects have frequently been accepted as natural or historically necessary. Yet these silences may in fact be produced through semantic and syntactic mechanisms internal to the text itself. The critical issue, therefore, is not simply who speaks, but who is granted the capacity to speak.

In recent years, discourse theory and poststructural approaches have enabled scholars to reread historical texts as systems of power and knowledge production. However, such theoretical perspectives often remain empirically under-supported and are rarely reinforced through computational methods. Consequently, historical silence is typically discussed at the level of interpretation rather than demonstrated as a structural pattern.

This study aims to address precisely this gap. By integrating poststructural theory with AI-based semantic analysis, it examines *Shiji* in order to identify relations of agency, silence, and power that are not explicitly articulated but are discursively embedded within the text. Artificial intelligence is not employed here as an interpretive authority, but as an analytic mechanism that renders discursive structures empirically visible.

The central research question guiding this study is as follows: How is silence produced in *Shiji*, and in whose interests does this silence operate? By addressing this question, the study argues that silence in historiography should be understood not as a contingent absence, but as an epistemic phenomenon structurally linked to power.

2. THEORETICAL FRAMEWORK

This study is theoretically grounded in a poststructuralist approach and examines the relationships between power, discourse, and silence in historical texts. Its theoretical foundation draws on three key conceptual frameworks developed by Michel Foucault, Gayatri Chakravorty Spivak, and Jacques Derrida. Together, these approaches allow historical texts to be read not merely as neutral accounts of past events, but as discursive systems that actively produce meaning, subject positions, and constraints.

The first theoretical pillar is based on Michel Foucault's concept of the relationship between power and discourse. According to Foucault, discourse is not only a medium for expressing knowledge but also a mechanism that constitutes and legitimizes power (Foucault, 2002; 1977; 1980). From this perspective, historical

texts do not simply describe events; rather, they define who is allowed to speak, which positions are rendered central, and which subjects are structurally consigned to silence. Historical works such as *Shiji* can therefore be understood as discursive apparatuses that organize visibility, authority, and exclusion.

The second theoretical pillar is grounded in Gayatri Chakravorty Spivak's concept of the subaltern. In Spivak's formulation, the subaltern subject is unable to speak not because of social origin alone, but because of the position assigned within discourse itself (Spivak, 1988). In this study, this framework is used to analyze how the discursive participation of envoys, captives, and non-central peoples is systematically constrained in historical narratives. Here, silence is not interpreted as a voluntary choice of the subject, but as a discursively produced condition.

The third theoretical foundation draws on Jacques Derrida's reflections on signification and meaning. According to Derrida, meaning is never fully present or immediately accessible; it is always deferred and mediated through signs (Derrida, 1976). This perspective enables an analysis of how symbols and signifiers— particularly those associated with loyalty and authority—delay the subject's voice and subordinate it to the discourse of central power. As a result, the subject is no longer the owner of meaning but becomes a carrier of signs.

Taken together, the theoretical framework of this study approaches historical texts not as inclusive or neutral sources, but as discursive systems that embody mechanisms of power, silence, and knowledge production. This framework provides the conceptual groundwork for the AI-based semantic analysis employed in subsequent sections and enables historical silence to be examined as an empirically analyzable phenomenon rather than a mere absence of speech.

3. LITERATURE REVIEW

Existing scholarship on Shiji and the intellectual legacy of Sima Qian has largely approached the text within the frameworks of historiography and narrative analysis. Classical studies have examined textual structure, authorial position, and the politics of memory (Durrant, 2017; Vankeerberghen, 2007; Horváth, 2010). However, these works rarely foreground the discursive marginalization of the Western Regions, envoys, and other peripheral subjects as a systematic analytical problem. More recent studies address ideological and historical ambivalence in Shiji (Lee, 2025), yet their analyses remain primarily confined to historical reconstruction rather than discursive mechanisms.

Theoretical literature provides a strong conceptual foundation for analyzing power–discourse relations. Foucault conceptualizes discourse as a mechanism that simultaneously produces knowledge and power (Foucault, 2002; 1977; 1980). Spivak frames the subaltern's inability to speak as the result of discursive positioning rather than social origin alone (Spivak, 1988). Derrida's theory of

deferred meaning explains how signification is mediated through symbols and absence (Derrida, 1976). Secondary interpretations of the dispositif have further elaborated its analytical potential (Callewaert, 2017; Raffnsøe, 2014; Pløger, 2023; 2008), yet their empirical application to East Asian historical texts remains limited (Crano, 2020).

Studies focusing on language and agency have identified grammatical and semantic mechanisms that constrain subjectivity (He, 2019), but the epistemic consequences of discursive silence are rarely connected to broader configurations of power. Cultural and literary research has explored center–periphery relations (Tang, 2000; Riemenschnitter, 2001), while regional historical studies have provided important empirical context (Ahmedov, 2023; Khudoyorov, N. M., and Ruziboev, S, 2025; Marsili, 2011). Nevertheless, these works generally do not aim to uncover discursive mechanisms at an analytic or structural level, nor do they employ computational methods.

Recent computational and AI-based approaches have opened new possibilities for historical analysis (OpenAI, 2025), yet such methods require careful theoretical grounding and methodological restraint. Studies that integrate theoretical discourse analysis, empirical textual evidence, and computational techniques in the examination of discourse, silence, and agency remain scarce. This study aims to address precisely this gap by combining poststructural theory with AI-based semantic analysis to examine historical silence as a structured and measurable phenomenon.

4. METHODOLOGY: AI-BASED SEMANTIC ANALYSIS

This study combines qualitative discourse analysis with computational, AI-based methods to examine historical texts at both discursive and structural levels. The methodology is not intended to automate interpretation; rather, it seeks to render relations of agency, silence, and power empirically visible within the textual structure.

The empirical corpus consists of *juan* 123 ("Dawan liezhuan") of Sima Qian's Shiji. This section was selected because it concentrates interactions among envoys, peoples of the Western Regions, and the imperial centre, thereby providing a focused textual space in which discursive silence and agency can be systematically examined. The text, written in Classical Chinese, was analysed using a single normalised version to ensure internal consistency. Prior to analysis, punctuation and quotation conventions were standardised, after which the text was tokenised. Key lexical units—such as verbs, subject designations, and central conceptual terms—were identified, while functional markers were recorded separately to account for the morphological characteristics of Classical Chinese. These steps constituted the structural preparation necessary for AI-assisted analysis.

Semantic analysis proceeded through three complementary procedures. Token frequency analysis was used to examine the distribution of subjects and verbs across active and passive constructions; co-occurrence analysis mapped semantic relationships among central concepts such as loyalty, authority, and envoyhood; and semantic clustering distinguished between high- and low-frequency units. Low-frequency but contextually salient units were interpreted as potential traces of discursive silence rather than as statistically insignificant data.

Artificial intelligence in this study functions as an analytical instrument rather than an autonomous interpretive subject. Its role is to assist in identifying latent structural patterns within the text. AI-generated outputs are not treated as conclusions in themselves but are interpreted theoretically within historiographical and cultural frameworks. Interpretive responsibility therefore remains entirely with the researcher.

AI-based semantic analysis entails several limitations. The high degree of contextual dependence in Classical Chinese may result in semantic loss during tokenisation and clustering, and frequency or co-occurrence patterns cannot fully represent discursive silence, serving instead as indicators of structural tendencies. The methodology is therefore interpretive rather than deterministic and does not aim to produce universal conclusions.

AI was employed solely as a support tool for detecting structural and semantic patterns. It was not used to generate historical arguments, interpretations, or conclusions. All analytical decisions and scholarly judgments remain the responsibility of the author. The limitations of AI—including potential bias, contextual reduction, and probabilistic error—are explicitly acknowledged, and AI-assisted results are treated as indicative rather than authoritative, with human interpretation retaining epistemic priority throughout the research process.

5. 使者之道 Shǐzhě zhī dào (The Way of the Envoy: The Discursive Nature of Diplomacy)

The following sentence by Sima Qian articulates a powerful semantic construction of imperial ideology: 張騫使西域，十餘年不忘漢節。 (*Zhāng Qiān shǐ Xīyù, shí yú nián bù wàng Hàn jié — Zhang Qian was sent as an envoy to the Western Regions and did not forget the Han emblem for more than ten years.*)

Here, the graph 使 (*shǐ*, "to send; to dispatch as an envoy") signifies more than a simple act of dispatch. It marks an active node within the imperial apparatus of knowledge and power. Within this discourse, Zhang Qian is no longer an individual historical figure but becomes a symbolic function—a mode through which the empire perceives and apprehends the external world. In Foucauldian terms, he operates as a sensory extension within the *dispositif* (apparatus) of imperial governance (Callewaert, 2017, pp. 29–52).

Similarly, the sign 節 (*jié*, "envoy emblem; symbol of loyalty") functions as a central semantic node expressing allegiance to sovereign authority. It elevates the envoy's speech, uttered in the name of the ruler, to the level of law and mandate (Crano, 2020). Even though Zhang Qian remains in captivity, his condition of **不忘漢節** (*bù wàng Hàn jié*, "not forgetting the Han emblem") positions him not as a speaking historical subject, but as a discursive apparatus that guarantees the continuity of imperial authority (Raffnsøe, 2014, pp. 278–298).

Sima Qian thus presents diplomacy as a political act through which the empire confirms itself by "seeing the Other," while simultaneously rendering the external world discursively silent (Thomas, 2008). Diplomacy, in this sense, is not merely a visit or a mission; it constitutes a central discursive mechanism of imperial domination (Pløger, 2008, pp. 57–70).

6. 俘與節之間 Fú yǔ jié zhī jiān (Between Captivity and Loyalty: The Paradox of Captivity and Allegiance)

Through the following sentence, Sima Qian articulates a complex discursive relationship between captivity and loyalty: 至匈奴，為單于所留，十年不得歸。 (*Zhì Xiōngnú, wéi Chányú suǒ liú, shí nián bù dé guī — Upon reaching the Xiongnu, he was detained by the Chanyu and was unable to return home for ten years.*)

This sentence contains two distinct layers of silence. The first is physical silence: Zhang Qian is a captive, and his speech—言 (*yán*, "speech")—is politically suspended; he cannot speak. The second is discursive silence: the historical narrative itself interprets this silence as an expression of loyalty.

Sima Qian presents **不得歸** (*bù dé guī*, "unable to return") not merely as a condition, but as a virtue. In this way, 沈默 (*chénmò*, "silence") functions as a marker of moral elevation. Within the semantic structure of the historical text, 沈默 (*chénmò*, "silence") and 忠 (*zhōng*, "loyalty") become closely aligned, nearly synonymous concepts.

At this point, Gayatri Spivak's well-known question becomes relevant: "Can the subaltern speak?" (Spivak, 1988, pp. 271–313). Zhang Qian appears here as a subaltern subject: he is inscribed within the language of the empire, yet unable to speak within it. What speaks through him is the 中心之聲 (*Zhōngxīn zhī shēng*, "voice of the center"). Silence here is not a political metaphor, but a syntactic property of the text itself.

Thus, Zhang Qian exists within imperial discourse, but not as a speaking subject; rather, the voice of the center speaks through him. As L. K. Jenco (2007) observes, historical texts present silence not as the subaltern's choice, but as a semantic structure produced by the center—not a voice, but an echo (Jenco, 2007, pp. 741–

755). In a similar vein, Andrea Riemenschnitter writes: "Loyal silence is not the echo of the subject, but the echo of the center. What matters here is not who speaks, but who is made to speak—this distinction is crucial" (Riemenschnitter, 2001, p. 23).

In modern Chinese literature, silence is likewise often interpreted as a form of heroism. Tang (2000) offers a critical analysis of this phenomenon, noting: "Silence is framed as heroism, yet this heroism serves not the subject, but the narrative of the state." Jin Liu (2013) characterizes silence as "strategic stillness", arguing through the example of Zhang Qian that silent heroes are not autonomous agents but figures positioned within state discourse (Liu, 2013, p. 87).

Indeed, women are frequently represented through silence; however, this is not voicelessness *per se*, but a condition of being placed outside the domain of language. In the context of peasant novels, Li (2018) writes: "The subaltern always speaks in the language of others; rather than telling their own story, they perform roles assigned to them" (Li, 2018, p. 13). Andolfatto (2019), analyzing utopian narratives in China, concludes that the hero who silently accepts subjugation becomes a symbol of loyalty—one that ultimately serves an unchanging structure of power.

7. 沈聲之構 CHÉN SHĒNG ZHĪ GÒU (THE STRUCTURE OF SILENT VOICE: THE SYNTACTIC ARCHITECTURE OF SILENCE)

The language of Sima Qian exhibits a deeply architectural character. Verbs such as 曰 (*yuē*, "said"), 至見 (*zhì jiàn*, "went to see"), and 還報 (*huán bào*, "returned and reported") are reserved exclusively for central subjects. By contrast, designations such as 匈奴 (*Xiōngnú*, "Xiongnu"), 大宛 (*Dàyuān*, "Dayuan"), and 月氏 (*Yuèzhī*, "Yuezhi") frequently appear in nominalized forms or within 所 (*suǒ*) passive constructions, producing agentless and passive sentences.

A clear example is provided by the sentence: 單于以為奇，留之。 (*Chányú yǐ wéi qí, liú zhī — The Chanyu considered him extraordinary and detained him.*) The semantic center of this sentence lies in 以為奇 (*yǐ wéi qí*, "to judge as extraordinary"). Here, 奇 (*qí*, "extraordinary") does not express subjective astonishment, but rather an imperial act of decoding. The clause 留之 (*liú zhī*, "detained him") marks the beginning of silence: the subject does not speak—he is held.

As He (2019) notes, in both Classical and modern Chinese, patient–subject constructions emerge through the syntactic restriction of agency. This structural weakness manifests discursively as silence. In Sima Qian's text, such imbalances in agent–patient relations constitute the phenomenon of 沈聲 (*chén shēng*, "silent voice").

The separation of verbs from agents, together with the nominalization or passivization of subjects, signals not only grammatical silence, but also a shift in the semantic center of discourse (He, 2019). In AI-based semantic analysis, these

constructions are identified as unidirectional power structures within agent–patient relations (OpenAI, 2025).

Empirically, this imbalance is visible in *Table 1*, which shows that 張騫 (*Zhāng Qiān*) is linked to active verbs twelve times, 匈奴 (*Xiōngnú*) appears as an agent nine times, while 大宛 (*Dàyuān*) occurs with active verbs only three times. This asymmetry is further visualized in *Figure 1*, where central tokens dominate the distribution of discursive agency. These quantitative disparities constitute an empirical expression of discursive silence. Thus, 沈聲 (*chén shēng*, "silent voice") functions not only as a semantic concept, but also as a grammatical form.

Sima Qian organizes speech according to a strict hierarchical distribution:

- Speaking subjects: the emperor — 天子 (*Tiānzǐ*, "Son of Heaven"); the envoy — 使者 (*shǐzhě*, "envoy")
- Subjects spoken about: peoples of the Western Regions — 西域諸國 (*Xīyù zhū guó*, "the states of the Western Regions")
- Silent entities: the perceptual landscape itself — 山川 (*shānchuān*, "mountains and rivers"), 民俗 (*mínsú*, "customs"), 土地 (*tǔdì*, "lands")

8. 節之形而上 JIÉ ZHĪ XÍNG ÉR SHÀNG (THE METAPHYSICAL SEMANTICS OF LOYALTY)

In classical Chinese political philosophy, the sign 節 (jié, "envoy emblem; bamboo credential") operates as a symbolic object positioned between 信 (xìn, "trust") and 命 (mìng, "command, mandate"). Sima Qian deploys this sign on two levels.

At the moral level, the expression 不忘漢節 (bù wàng Hàn jié, "not forgetting the Han emblem") frames the envoy as a figure of personal sacrifice and ethical commitment. Loyalty here is not an internal virtue of the subject, but an obligation expressed through fidelity to a symbol.

At the epistemic level, 節 (jié) functions as the written trace of imperial authority—a sign through which the empire inscribes its mandate into discourse. In this sense, 節 becomes the material body of language itself.

AI-based semantic analysis demonstrates that 節 (jié) co-occurs seventeen times with 漢 (Hàn, "imperial center") and eleven times with 使 (shǐ, "to dispatch as envoy"), while it does not co-occur at all with 大宛 (Dàyuān). As shown in Table 1 and illustrated in Figure 1, this distribution positions 節 as a central semantic hub within the text (OpenAI, 2025). Historical loyalty is thus structurally tied to the center, while "others" remain outside this semantic network and are rendered discursively silent.

In this sense, 節 (jié) operates as a semantic mechanism that centralizes silence. Philosophically, this aligns with Derrida's concept of différance (Derrida, 1976, p. 61): 節 does not affirm presence directly, but rather defers absence. The envoy carries the emblem, yet the emblem simultaneously negates his personal voice. He becomes a bearer of the sign, but never the owner of its meaning.

9. 俘之倫理 Fú zhī lúnlǐ (THE ETHICAL SEMANTICS OF CAPTIVITY)

Within Han-dynasty historiography, the captivity and 沈默 (chénmò, "silence") of 張騫 (Zhāng Qiān) are not framed as mere passivity, but as a mechanism through which an ethical position is constructed. Silence does not indicate the absence of voice; rather, it operates as a deferred moral gaze mediated through writing. In this perspective, textual gaps and omissions function not as deficiencies of the record but as active components of historiographical meaning, shaping the ethical horizon within which the historical subject is perceived.

As Haraway argues, "situated knowledge"—knowledge rooted in context—appears in Sima Qian's text as "situated silence": silence that is conditioned and grounded by its context (Haraway, 1988, pp. 575–599). These unwritten truths become the primary source of ethical force, indirectly yet decisively shaping the position of the historical subject.

10. 無聲之聲 Wú shēng zhī shēng (THE ONTOLOGY OF THE SILENT VOICE)

Through the following sentence, Sima Qian articulates a complex ontological relationship between silence and writing: **乃得脱歸，具以西域之事上書。** (*Nǎi dé tuō guī, jù yǐ Xīyù zhī shì shàng shū — He finally escaped, returned home, and submitted a written report to the emperor concerning the Western Regions.*)

Two processes occur simultaneously in this sentence. The first, **脱歸** (*tuō guī*, "escape and return"), denotes physical liberation. The second, **上書** (*shàng shū*, "to submit a written memorial"), marks discursive re-entry.

For Sima Qian, freedom materializes only through writing. The envoy exits a state of silence and re-enters the language of the center. He can now "speak," yet his speech remains written for the center and in the name of the center. It is at this point that **無聲之聲** (*wú shēng zhī shēng*, "the silent voice") acquires ontological significance: it has not disappeared; rather, it exists metonymically, living through writing.

From the perspective of AI-based semantic analysis, this constitutes a crucial indicator: silence is not a zero frequency, but a low-probability mode of existence (OpenAI, 2025). In historical writing, **無聲之聲** (*wú shēng zhī shēng*, "the voice of the voiceless") can be detected even at the level of token frequency.

When Sima Qian writes about particular peoples, their own speech is absent, yet their 名 (*míng*, "names") are preserved. These names function as the statistical traces of silence, that is, as tokens. An AI semantic model can interpret such traces as latent variables, or hidden dimensions of meaning; through these dimensions, historical silence becomes algorithmically audible (OpenAI, 2025). In this way, Sima Qian's historiography advances a proto-semantic logic in which absence itself becomes a signal, and silence operates as data.

11. Conclusion

This study offers a rereading of Sima Qian's *Shiji* not as a conventional historical narrative but as a discursive system structured by mechanisms of power, silence, and knowledge production. By integrating poststructural theory (Foucault, Spivak, Derrida) with AI-assisted semantic analysis, it identifies "silent voices" that are not explicitly articulated yet remain structurally embedded within the text. The silence surrounding envoys, peoples of the Western Regions, and other peripheral actors is shown to be neither incidental nor merely narrative omission; rather, it emerges as a semantic and syntactic effect of imperial discourse. Imbalances in agent–patient constructions, the distribution of active and passive verbs, and semantic clustering around key concepts such as 節 (*jié*, envoy emblem/marker of loyalty), 忠 (*zhōng*, loyalty), and 漢 (*Hàn*, imperial centre) reveal a patterned dominance of central authority. Silence thus operates as an active discursive mechanism that reinforces imperial power rather than as a simple absence of voice.

Within this framework, AI functions not as an autonomous interpretive authority but as an analytical instrument for detecting latent structural patterns in historical writing. Its use enables silence to be examined beyond purely qualitative observation, rendering low-frequency yet contextually significant units empirically visible within the textual system. Subjects traditionally characterised as "voiceless" therefore appear not as null presences but as statistically marginalised yet meaningful components of the discourse. In theoretical terms, the study reconceptualises historical silence as a structural phenomenon and extends poststructural approaches into the analysis of classical historiography. Methodologically, it demonstrates how close reading and AI-assisted semantic analysis can be combined in a critical and responsible manner, with computational tools supporting—rather than replacing—human interpretation in the humanities.

BIBLIOGRAPHY

Ahmedov, A. P. (2023) *O'zbekiston eng qadimgi va qadimgi davrlarda (Paleolit davridan milodiy V asrgacha)*, Vol. 3. O'zbekiston: O'zbekiston NMIU.

Andolfatto, L. (2019) *Hundred days' literature: Chinese utopian fiction at the end of empire, 1902–1910*, *East Asian Comparative Literature and Culture*, Vol. 11. Leiden: Brill. doi: 10.1163/9789004398856.

Boltz, W. G. (1994) *The origin and early development of the Chinese writing system*, *American Oriental Series* 78. New Haven, CT: American Oriental Society.

Callewaert, S. (2017) 'Foucault's concept of dispositif', *Praktiske Grunde: Nordisk tidsskrift for kultur- og samfundsvidenskab*, 1–2, pp. 29–52.

Derrida, J. (1976) *Of grammatology*. Translated by G. C. Spivak. Baltimore, MD: Johns Hopkins University Press.

Durrant, S. V. (2017) 'The letter to Ren An and Sima Qian's legacy', *Early China*, 40, pp. 1–25. doi: 10.1017/eac.2017.6.

Fairclough, N. (1995) *Critical discourse analysis: The critical study of language*. London: Longman.

Foucault, M. (1977) *Discipline and punish: The birth of the prison*. New York: Pantheon Books.

Foucault, M. (1980) *Power/knowledge: Selected interviews and other writings, 1972–1977*. New York: Pantheon Books.

Foucault, M. (2002) *The archaeology of knowledge*. London: Routledge. doi: 10.4324/9780203604168.

Haraway, D. (1988) 'Situated knowledges: The science question in feminism and the privilege of partial perspective', *Feminist Studies*, 14(3), pp. 575–599. doi: 10.2307/3178066.

He, X. (2019) *Patient-subject constructions in Mandarin Chinese: Syntax, semantics, discourse*, *Studies in Chinese Language and Discourse* 12. Amsterdam: John Benjamins. doi: 10.1075/scld.12.

Horváth, I. (2010) 'Sima Qian and the Shiji: A lesson in historiography by a contemporary of the Xiongnu', in Geybullayeva, P. R. (ed.) *Stereotypes in literatures and cultures*, pp. 117–129. Peter Lang.

Jenco, L. K. (2007) 'What does heaven ever say? A methods-centered approach to cross-cultural engagement', *American Political Science Review*, 101(4), pp. 741–755. doi: 10.1017/S000305540707043X.

Khudoyorov, N. M. and Ruziboev, S. (2025) 'Sima Qian's Shiji work: Historical truth or ideological construction?', *Journal of Modern Educational Achievements*, 3, pp. 31–44.

Lee, J. (2025) 'Duality of historian Sima Qian: Reconstructing the Yao and Shun narratives as ideological and historical writings in the Shiji', *Journal of Asian Historical Studies*, 6, pp. 171–315. doi: 10.17856/JAHS.2025.6.171.315.

Li, T. (2018) 'Novels of Zhao Shuli and Sun Li: Chronicles of new peasantry', in McDougall, K. and Fong, A. (eds.) *Routledge handbook of modern Chinese literature*, pp. 311–323. London: Routledge.

Liu, J. (2013) *Signifying the local: Media productions rendered in local languages in mainland China in the new millennium*. Leiden: Brill. doi: 10.1163/9789004259027.

Meng, Y. and Dai, J. (2023) *Emerging from the horizon of history: Modern Chinese women's literature, 1917–1949*. Singapore: Palgrave Macmillan.

OpenAI (2025) *ChatGPT (GPT-4/GPT-4o)* [large language model]. Available at: https://chat.openai.com (Accessed: 14 January 2026).

Pløger, J. (2008) 'Foucault's dispositif and the city', *Planning Theory*, 7(1), pp. 51–70. doi: 10.1177/1473095207085665.

Raffnsøe, S., Gudmand-Høyer, M. and Thaning, M. (2014) 'Foucault's dispositive: The perspicacity of dispositive analytics in organizational research', *Organization*, 23(2), pp. 272–290. doi: 10.1177/1350508414549885.

Riemenschnitter, A. (2001) 'Mother China myths in twentieth-century literary narratives', in *Asian nationalism in an age of globalization*, pp. 23–38. London: Routledge.

Sima Qian 司馬遷 (2025) *Shi ji 史記*, juan 123 'Dawan liezhuan 大宛列傳'. Beijing: Zhonghua guji tushuguan digital edition. Available at: http://www.gmzm.org/?gujitushu/shiji.html (Accessed: 14 January 2026).

Spivak, G. C. (1988) 'Can the subaltern speak?', in Nelson, C. and Grossberg, L. (eds.) *Marxism and the interpretation of culture*, pp. 271–313. Urbana, IL: University of Illinois Press.

Tang, X. (2000) *Chinese modern: The heroic and the quotidian*. Durham: Duke University Press.

Thomas, K. D. (2008) 'The history of others: Foreign peoples in early Chinese historiography'. Available at: https://www.ias.edu/ideas/2008/di-cosmo-historiography (Accessed: 14 January 2026).

van Dijk, T. A. (1998) *Ideology: A multidisciplinary approach*. London: Sage.

Vankeerberghen, G. (2007) 'The tables (Biao) in Sima Qian's Shiji: Rhetoric and remembrance', in *The warp and the weft*, pp. 39–58. Leiden: Brill. doi: 10.1163/ej.9789004160637.i-772.39.

Yang, L. (2023) *Building blocks of Chinese historiography: A narratological analysis of the Shiji*. PhD thesis. University of Pennsylvania. Available at: https://repository.upenn.edu (Accessed: 14 January 2026).

GLOSSARY OF KEY TERMS

- 節 (jié): Envoy emblem; bamboo credential symbolizing loyalty and delegated authority. In Shiji, jié functions as a discursive marker through which imperial mandate is materially inscribed and centralized.

- 沈聲 (chén shēng): "Silent voice." A syntactic–semantic condition in which agency is structurally constrained through passive constructions and agent–patient asymmetry, producing silence within discourse rather than mere absence of speech.

- 無聲之聲 (wú shēng zhī shēng): "The voice of silence." An ontological concept denoting the continued existence of subaltern presence through writing, naming, and low-frequency textual traces rather than direct speech.

- 中心之聲 (zhōngxīn zhī shēng): "Voice of the center." The authoritative discourse of imperial power that speaks through peripheral subjects, replacing their individual agency with centralized narration.

- 使 (shǐ): To dispatch; to send as an envoy. Beyond a verb of motion, shǐ marks incorporation into the imperial apparatus of knowledge, surveillance, and authority.

- 忠 (zhōng): Loyalty. In Shiji, zhōng is not an internal moral quality but a discursively produced virtue aligned with obedience to imperial symbols and mandate.

- 所 (suǒ): A grammatical marker forming passive or agentless constructions in Classical Chinese. In this study, suǒ-constructions are treated as key syntactic mechanisms of discursive silence.

- 名 (míng): Name. The preservation of names without accompanying speech functions as a statistical and semantic trace of silenced subjects within historical writing.

FIGURES
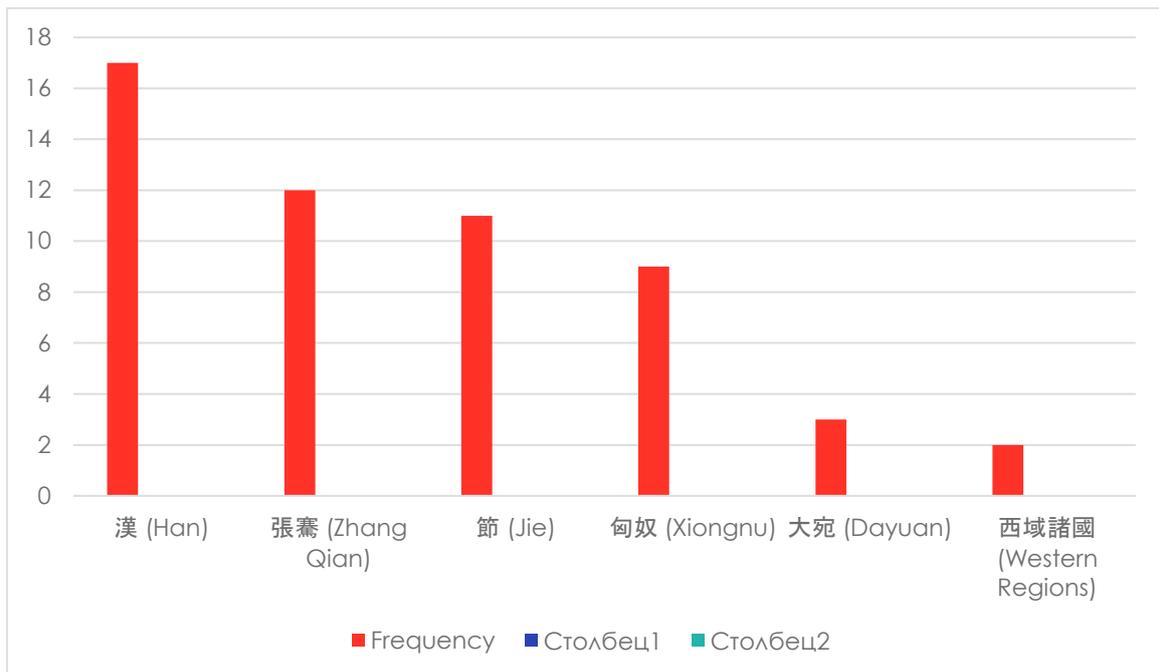
Figure 1. Distribution of discursive agency in Shiji, Juan 123 (Dawan liezhuan).

TABLES

Entity / Subject	Active Verb Occurrences	Passive / Agentless Constructions	Discursive Role
Zhang Qian (張騫)	12	2	Central envoy / authorized agent
Xiongnu (匈奴)	9	3	Semi-active external power
Dayuan (大宛)	3	7	Peripheral, largely silenced subject
Yuezhi (月氏)	2	6	Peripheral, nominal presence
Han / Imperial Center (漢 / 天子)	15	0	Absolute discursive center

Table 1. Distribution of discursive agency in Shiji, Juan 123 (Dawan liezhuan).

Divulgación vs Vulgarización: el reto de la creación de contenidos en un mundo saturado

Public Outreach vs. Oversimplification:
The Challenge of Content Creation in a Saturated World

Mario Agudo Villanueva

Codirector de Karanos
marioagudo@gmail.com
0000-0003-3364-231X

La Historia se construye mediante un método basado en hipótesis formuladas a partir de la información que proporcionan las fuentes literarias o los hallazgos arqueológicos y se sirve de publicaciones académicas, congresos o de formatos de alta divulgación para darse a conocer. Sin embargo, estos canales suelen tener un alcance restringido a círculos minoritarios, a pesar del esfuerzo de difusión de los últimos años. Paradójicamente, los datos demuestran que existe cierto interés social, aunque ni mucho menos generalizado, en esta disciplina. Según la segunda oleada de 2025 del Estudio General de Medios (EGM), elaborado por la Asociación de Investigación de Medios de Comunicación (AIMC), entre las diez revistas temáticas mensuales más leídas hay cuatro que tienen contenidos históricos o relacionados con el patrimonio (cuadro 1). Sin embargo, la impresión generalizada entre los profesionales es que el acercamiento del público generalista a sus especialidades se produce a través de tópicos o falsedades. Analizar las razones de esta aparente distorsión es una tarea compleja, cuyo alcance sería mucho mayor del que podría abarcarse en este artículo.

1. UN CASO PRÁCTICO: EL CLAUSTRO DEL *MAS DEL VENT*

En diciembre de 2010, la revista *Románico* (de la cual era director en aquel momento), dedicada a la alta divulgación, publicó en exclusiva un artículo del profesor Gerardo Boto Varela, de la Universitat de Girona, en el que se anunciaba el descubrimiento de un claustro medieval en la finca del Mas del Vent, en Palamós. A pesar de la espectacularidad del hallazgo, la noticia no se propagó de forma masiva hasta que el diario *El País* se hiciera eco de ella después de unas jornadas organizadas por la Universitat de Barcelona en mayo de 2012 bajo el título *Arte fugitivo*, en las que el propio autor daba cuenta de sus últimas investigaciones. Es decir, mediaron casi dos años entre el primer artículo y su salto a los medios de comunicación, pero, a excepción de algunos círculos restringidos, no se produjo una

respuesta social significativa hasta que la cuestión irrumpió en la agenda mediática. Las redes sociales estallaron con la noticia del descubrimiento del claustro, mientras habían guardado silencio cuando era conocido únicamente en círculos académicos. Ni el mundo profesional, ni una publicación especializada, fueron capaces de introducir en la agenda mediática un tema tan atractivo como el descubrimiento de una obra de arte inédita[1].

Uno de los pioneros de los estudios académicos sobre los efectos sociales de los medios de comunicación, Walter Lippmann, señalaba que el análisis de la opinión pública debía comenzar por el reconocimiento de la relación entre la acción, la imagen de esa acción y la respuesta que el hombre generaba ante la imagen que se había forjado de tal acción[2]. En este sentido, según el autor norteamericano, son los medios de comunicación los que generan las imágenes con las que interpretamos lo que acontece en todo el mundo. En la medida en la que este universo está fuera de nuestro alcance, algo que ocurre con frecuencia, nuestras mentes pueden generar imágenes diferentes al mundo real[3]. Podríamos decir que los medios proyectan sombras a través de los que interpretamos nuestro entorno, como una suerte de mito de la caverna contemporáneo.

La clave estaría, por tanto, en introducir temas en la agenda mediática, pero esto no es una tarea sencilla. Los motivos que explican el escaso protagonismo de las informaciones relacionadas con la historia o el patrimonio son profundos: los medios son empresas informativas que necesitan ingresos procedentes, en su mayor parte, de la publicidad, lo que supone que tengan que buscar el rédito económico de cada centímetro cuadrado, línea, minuto o segundo que publiquen o emitan. Lo más noticiable será aquello que consiga una mayor audiencia, luego estamos ante un problema más estructural que coyuntural y que no solo afecta al mundo de la difusión. Esta situación no deja de ser preocupante, máxime si tenemos en cuenta planteamientos teóricos como los de Walter Lippmann o Maxwell McCombs[4], que muestran de forma meridiana cómo la sociedad se preocupa y genera opinión solo en base a los temas que se introducen en la agenda mediática. La invisibilidad en los medios condena a una inexistencia en la práctica, incluso para la administración pública, pues también condicionan la agenda política[5].

2. UN NUEVO PARADIGMA

Semejante panorama es propio de un modelo en el que los principales canales de comunicación eran controlados por los medios que hoy consideramos

---

[1] Sobre el caso del claustro de Palamós, véase Agudo Villanueva (2018).
[2] Lippmann (1922, p. 16).
[3] Lippmann (1922, p. 29).
[4] Sobre la construcción mediática de las imágenes del mundo y su jerarquización temática, véanse Lippmann (1922) y McCombs (1996).
[5] Para una síntesis clásica del desarrollo de la investigación en agenda-setting, véase Rogers and Dearing (1988, pp. 555–594).

tradicionales, monopolizadores del mensaje masivo. Sin embargo, en el siglo XXI se ha gestado un nuevo paradigma provocado por la irrupción de las redes sociales. La posibilidad de conexión de los públicos entre sí ha propiciado un modelo de comunicación nodal, una situación en la que el canal está abierto para toda aquella institución, grupo o persona que quiera dirigir sus mensajes a una comunidad determinada. Estamos, por tanto, ante una alternativa de enorme potencial para afrontar la tarea de la divulgación, pero que complica todavía más nuestra propuesta de análisis. Pero este paradigma también plantea importantes retos. El incremento de la accesibilidad a los canales de comunicación ha multiplicado exponencialmente la exposición a estímulos informativos, fenómeno que Alfons Cornella bautizó hace ya unos años, de forma lúcida, como infoxicación[6]. Asistimos a la paradoja de que cuanta más información tenemos, cuando más conectados nos encontramos, más desinformados estamos. Es difícil no recordar aquellos proféticos versos de Thomas Stearns Eliot en los que se preguntaba, allá por el año 1922, dónde estaban la sabiduría y el conocimiento que se habían perdido en la información[7].

Los medios de comunicación han tenido la inmediatez como referente a lo largo de su historia. Cada avance tecnológico permitía acercarse al objetivo del tiempo real, algo que ya hemos conseguido gracias a la combinación de smartphones, mejoras de conectividad y redes sociales. Pero la sucesión de informaciones transmitidas sin solución de continuidad impide que se investiguen de una manera adecuada, por lo que se pierde profundidad en el tratamiento[8]. Se producen contenidos más superficiales, pero también se consumen de manera menos crítica. Como agravante, la imagen ha derrotado al texto y las opiniones se forjan sobre simples apariencias. Por otro lado, la permanencia de los acontecimientos en la agenda es más efímera. La actualidad de un mundo global hace que las informaciones se agolpen, entren y salgan de nuestras pantallas con una gran celeridad. La consecuencia más evidente es que se pierde la memoria. Estamos ante otra paradoja: en el mundo del boca a boca por escrito y viral, donde podemos rastrear el testimonio de cualquier persona conectada a las redes, la memoria se diluye ante la avalancha de acontecimientos. Si agitamos la coctelera de la sobreinformación, la superficialidad y la falta de memoria nos encontramos con uno de los principales problemas de la divulgación: la falta de rigor[9].

Conviene, llegados a este punto, realizar una matización terminológica. Señala la Real Academia de la Lengua Española que divulgar es "publicar, extender, poner al alcance del público algo"[10]. Sin embargo, para hacerlo se suele rebajar el tono,

---

[6] Sobre el concepto de *infoxicación* y la sobrecarga informativa en entornos digitales, véase Cornella (2003).

[7] Para la metáfora cultural de la fragmentación y el colapso del sentido en la modernidad, puede evocarse Eliot (1922/2009).

[8] Sobre el uso de redes sociales como herramienta de mediación patrimonial y divulgativa, véase Agudo Villanueva (2017, pp. 112–119).

[9] Sobre la construcción de la opinión pública en entornos digitales y la metáfora de la "caverna mediática", véanse Agudo Villanueva (2016a; 2016b).

[10] Para la definición lexicográfica de "divulgar", véase RAE (s. f.).

distorsionar el mensaje, caer en lo anecdótico, sobre la falsa concepción de que así será entendido por un porcentaje más amplio de la audiencia y, por tanto, tendrá más impacto. Cuando caemos en esta tentación corremos el riesgo de vulgarizar el contenido, es decir, acudiendo de nuevo a la Real Academia, "exponer una ciencia, o una materia técnica cualquiera, en forma fácilmente asequible al vulgo"[11].

3. OTROS ACTORES: LA FICCIÓN

Es evidente, por tanto, que entre el conocimiento generado en el ámbito académico y la sociedad se tejen una serie de conexiones sobre las que debemos reflexionar como punto de partida, pues resultan determinantes en el análisis de la imagen que se proyecta de los acontecimientos históricos y sus protagonistas. Esta labor de intermediación podría sintetizarse en un modelo de propagación (como el que se muestra en el *cuadro 2*). En muchas ocasiones, más de las deseables, la fuente primaria no es el mundo académico, donde se genera el conocimiento de forma metodológica, sino fuentes de diversa índole en las que no vamos a profundizar en este artículo, pero cuyo impacto no es, en absoluto, desdeñable. Una de las más importantes, por su extraordinaria capacidad de penetración, es la ficción.

Cada vez que se estrena una nueva película o serie televisiva de temática histórica se generan interminables debates públicos sobre su veracidad, en especial, en las redes sociales. Esta dinámica es semejante a la que acompaña a los lanzamientos de novelas, aunque en los libros encontramos un agravante. Algunos autores han adquirido la costumbre de citar una extensa bibliografía, que a veces sobrepasa la decena de páginas, como justificación de un prolijo proceso de documentación que, si bien resulta meritorio, no deja de insinuar que el producto ante el que nos encontramos goza de una mayor rigurosidad. Sin embargo, enumerar una larga retahíla de investigaciones sobre la materia que sirve de marco para la obra de ficción no convierte a la novela en un ensayo. La confusión aumenta cuando los medios de comunicación se hacen eco de la novedad mediante entrevistas en las que se considera al escritor una fuente de autoridad en el período histórico correspondiente, lo que da pie a ejercicios de presentismo encaminados a captar la atención de los lectores. Mientras el relato histórico se construye mediante una rigurosa metodología, la ficción se basa en un proceso creativo ambientado en un período histórico en el que el autor debe optar por una trama concreta y una determinada caracterización de los personajes. En tanto que la Historia es un relato abierto que tiene el objetivo de generar conocimiento; la ficción propone, necesariamente, un relato cerrado con el que se pretende entretener, para lo que se ve en la obligación de completar las comprensibles lagunas epistemológicas de nuestro pasado tirando de imaginación.

---

[11] Para la definición lexicográfica de "vulgarizar", véase RAE (s. f.).

La ficción tiene un alcance determinante, pues la industria cultural se sirve de los medios de comunicación de masas, con los que a veces mantiene intereses empresariales comunes, para maximizar su impacto. Una situación que las redes sociales, a pesar de haber revolucionado el modelo tradicional de comunicación, no han conseguido paliar todavía, si es que están en disposición de hacerlo. Podemos afirmar que el imaginario colectivo contemporáneo está compuesto de imágenes vistas, por lo que los medios adquieren un indiscutible protagonismo en la transmisión de contenidos culturales[12]. A ello debemos añadir el potencial de otros componentes de la industria del ocio, como los videojuegos, que terminan por apuntalar un complejo conglomerado que ejerce una notable influencia en la manera en que percibimos el relato histórico[13].

4. Realidad, verdad racional y verdad factual

Si los medios de comunicación tejen un conjunto de relatos de la realidad a partir de la selección de hechos debemos centrarnos, por tanto, en qué entendemos por realidad. No es el propósito de este texto repasar el estado de la cuestión filosófica sobre este imbricado concepto. Basta para nuestro fin recurrir a la definición que de ella se proporciona en el diccionario de la Real Academia Española, donde se explica que realidad es la "existencia real y efectiva de algo" o, en segunda acepción, "la verdad, lo que ocurre verdaderamente"[14]. En sentido aristotélico podríamos asegurar que realidad es lo que es y no es lo que no es, pues "decir del Ser que no es o del No Ser que es, es falso, mientras que decir del Ser que es y del No Ser que no es, es verdadero"[15]. Hannah Arendt distinguía entre dos tipos de verdad: la racional, que es la verdad científica, cuya antítesis sería el error o la ignorancia, y la factual, relativa a los hechos, cuya antagonista es la mentira[16]. Sin embargo, acceder a la verdad, en ambos casos, no es tarea sencilla, en palabras de Heráclito: "la verdadera naturaleza gusta de ocultarse"[17].

Como hemos visto, el modelo lineal de comunicación, en el que los medios informativos disponían del monopolio del mensaje se ha transformado en un nuevo paradigma en el que la emisión ha dado paso a la conversación. La función de intermediación de los medios se desvanece ante la aparición de "creadores de contenido" y "medios alternativos", que enarbolan la bandera de la libertad de expresión para realizar una actividad paralela que, en el mejor de los casos, no se guía por criterios profesionales y, en el peor, responde a intereses partidistas. No todo en las redes es negativo, ni mucho menos, pero su integración en el ecosistema informativo ha generado una importante convulsión cuya consecuencia más evidente es el incremento exponencial del contenido que desfila ante nosotros. Se

---

[12] Para una reflexión temprana sobre la relación entre cine e historia, véase Montero and Paz (1997, pp. 7–9).

[13] Sobre la influencia de las imágenes digitales en lamemoria histórica, véase Venegas Ramos (2023).

[14] Para las acepciones lexicográficas del término "realidad", véase RAE (s. f.).

[15] Aristóteles (1011b 25).

[16] Sobre la tensión entre verdad factual y discurso político, véase Arendt (2017, pp. 22–23).

[17] Heráclito, fr. 8 Marc. (B 123 DK) en Bernabé (2019, p. 176).

ha abierto una guerra atroz por la captación de la atención, lo que no ha hecho más que profundizar en la crisis de los contenidos[18]. Se da la paradoja de que cuanta más información tenemos, cuando más conectados nos encontramos, más desinformados estamos. En efecto, la acumulación de contenidos no significa que estemos más cerca de la verdad, pues llega un punto en el que el exceso se convierte en inabordable, inclasificable e inútil[19]. Harry Frankfurt ha publicado recientemente dos ensayos a propósito de cómo la proliferación de personas dedicadas a transmitir información en este nuevo ecosistema, a las que denomina charlatanes, son en realidad impostores que pretenden manipular las opiniones y las actividades de la audiencia[20].

5. POSVERDAD, *FAKE NEWS* Y DESINFORMACIÓN

El consumo de medios de comunicación siempre ha adolecido de un notable sesgo de confirmación, la tendencia a elegir voces que refuercen nuestra manera de pensar, que nos hagan sentir miembros de una determinada comunidad. Estamos dispuestos a creer aquello que refuerza nuestra identidad. La irrupción de las redes sociales no ha hecho más que reforzar esa tendencia, pues dentro de la amplia gama de oferta informativa presente en este nuevo ecosistema, cada usuario puede forjarse su particular "cámara de eco", lo que acaba por generar la sensación de que nuestros intereses son los únicos que existen o, al menos, los más legítimos. Estas aventuras identitarias, sean de la naturaleza que sean, no hacen más que minar el sentido de comunidad. Como agravante, la pugna por la atención ha desencadenado profundos cambios en los criterios de noticiabilidad. Se recurre a lo conflictivo, lo escandaloso, lo provocador o lo personal, generalmente porque se trata de acontecimientos que generan emociones, que no se dirigen a la esfera racional del ser humano, sino a la sentimental. El resultado es una creciente polarización en la que parecemos más dispuestos que nunca a creer algo porque, sencillamente, es una cuestión de "los nuestros" contra "los otros"[21].

Esta coyuntura es tierra abonada para la posverdad, que podríamos considerar como el desprecio más profundo de la verdad, cuando el sentimiento importa mucho más que los hechos. Un mundo de opiniones fundamentadas en una suerte de supremacía ideológica en la que podemos desafiar la realidad sin rendir cuentas, amparados en esa comunidad dispuesta a aceptar cualquier información que refuerce sus creencias[22]. La posverdad tiene viejas raíces, pues no deja de ser una versión cínica del relativismo más extremo, que podría encontrar su epítome en el

---

18 Sobre la crisis estructural del periodismo contemporáneo y sus implicaciones en la esfera pública, véase León Gross (2024, pp. 202–203).

19 Para el análisis de la transformación de los marcos informativos y su impacto en la credibilidad mediática, véase León Gross (2024, pp. 194–195).

20 Sobre la distinción conceptual entre verdad, mentira y *bullshit*, véanse Frankfurt (2005 y 2006).

21 Para la relación entre degradación del discurso público y pérdida de mediación periodística, véase León Gross (2024, p. 108).

22 Para una caracterización filosófica de la posverdad, véase McIntyre (2023, pp. 34–43).

viejo aforismo de Protágoras que situaba al hombre como medida de todas las cosas[23]. Pero su antecedente más reciente es el posmodernismo, una postura filosófica que parte de la negación total de la verdad objetiva, pues asegura que cualquier posicionamiento sobre ella no es nada más que un reflejo de la ideología política de la persona que lo pronuncia[24]. Cuando el único criterio de veracidad son las opiniones, las creencias o las emociones, no nos encontramos con ciudadanos informados, sino movilizados[25]. Parecen muy lejanas aquellas palabras de Charles Prestwich Scott en 1921, cuando The Guardian era The Manchester Guardian, en las que afirmaba que "comentar es libre, pero los hechos son sagrados"[26]. El pionero editor marcaba una clara línea roja entre información y opinión, pero casi un siglo después esos límites parecen más difusos que nunca.

No debemos confundir la posverdad con otro mal de nuestro tiempo, que siempre ha existido pero que ha encontrado en este nuevo ecosistema un lugar en el que campar a sus anchas: la mentira. Mientras que la posverdad no reconoce ningún valor a la verdad, el mentiroso sí, pero la adultera con una intención determinada. Las noticias falsas propagadas con algún fin reciben en castellano el nombre de "bulos", aunque su uso en la modernidad se ha popularizado con el término anglosajón fake news. Es preferible, sin embargo, el concepto de "noticia falseada", tal como recomienda Fundeu, ya que la falsedad puede ser, a veces, involuntaria, pero el que falsea lo hace con un propósito[27]. El falseamiento de la realidad es la piedra angular de otro concepto decisivo: la desinformación, la difusión deliberada de información falsa o engañosa con el propósito de manipular, confundir o influir en la percepción pública y las decisiones de las personas. Las nuevas tecnologías ponen a nuestra disposición herramientas que han afinado la ya de por sí extraordinaria capacidad del ser humano para mentir. Hoy en día podemos mentir mejor y, sobre todo, propagar la falsedad con más facilidad. Es indispensable que desde el ámbito profesional y desde las instituciones se trabaje de forma activa en combatir la mala praxis, pero buena parte de la responsabilidad de que este tipo de contenidos se difundan es nuestra. Los ciudadanos debemos realizar un esfuerzo consciente para no contribuir en la propagación de este tipo de informaciones, debemos aguzar el pensamiento crítico y no ser cómplices de la mentira recordando que tenemos el derecho constitucional de recibir información veraz.

---

[23] Platón (*Teeteto* 152a).
[24] Sobre la consolidación del discurso posverdadero en el espacio público, véase McIntyre (2023, p. 139).
[25] Para el impacto de este fenómeno en la práctica periodística contemporánea, véase León Gross (2024, p. 215).
[26] Sobre la ética periodística clásica y la función normativa de la prensa, véase *The Guardian* (s. f.).
[27] Para la definición y uso del término "fake news" en el ámbito hispanohablante, véase FundéuRAE (s. f.).

BIBLIOGRAFÍA

Agudo Villanueva, M. (2016a) 'La caverna mediática, las redes sociales y la opinión pública'. Disponible en: http://www.tecnonews.info (Consulta: 11 mayo 2017).

Agudo Villanueva, M. (2016b) 'Arqueonet – hacia una propuesta de valor en divulgación'. Disponible en: http://www.mediterraneoantiguo.com (Consulta: 11 mayo 2017).

Agudo Villanueva, M. (2017) 'Patrimonio y redes sociales: una alternativa para la divulgación', en Frías Castillejo, C. y López Mira, J. A. (eds.) *Actas de las II Jornadas de Museos y Colecciones Museográficas Permanentes de la Comunidad Valenciana. Nuevas tecnologías aplicadas a la gestión turística del patrimonio*. Alicante, pp. 112–119.

Agudo Villanueva, M. (2018) 'Periodismo, patrimonio y agenda setting: el claustro de Palamós como caso de estudio de impacto mediático', en Boto Varela, G. (ed.) *Salamanca–Ciudad Lineal–Palamós. Las arcadas claustrales de Mas del Vent*. Salamanca, pp. 27–40.

Arendt, H. (2017) *Verdad y mentira en la política*. Barcelona: Página Indómita.

Bernabé, A. (2019) *Fragmentos presocráticos*. Madrid: Abada.

Cornella, A. (2003) *Infoxicación: buscando un orden en la información*. Barcelona: Infonomia.

Eliot, T. S. (2009) *La tierra baldía*. Trad. J. L. Palomares. Madrid: Cátedra.

Frankfurt, H. (2005) *On Bullshit*. Princeton: Princeton University Press.

Frankfurt, H. (2006) *On Truth*. New York: Knopf.

FundéuRAE (2017) 'Noticia falsa, falseada, fake news'. Disponible en: https://www.fundeu.es/recomendacion/noticia-falsa-falseada-fake-news/ (Consulta: 7 enero 2025).

León Gross, T. (2024) *La muerte del periodismo*. Barcelona: Deusto.

Lippmann, W. (1922) *Public Opinion*. New York: Harcourt, Brace and Company.

McCombs, M. (1996) 'Influencia de las noticias sobre nuestras imágenes del mundo', en Bryant, J. y Zillmann, D. (eds.) *Los efectos de los medios de comunicación. Investigaciones y teorías*. Barcelona: Paidós, pp. 13–34.

McIntyre, L. (2023) *Posverdad*. Madrid: Cátedra.

Montero, J. y Paz, M. A. (eds.) (1997) *La historia que el cine nos cuenta. El mundo de la posguerra, 1945–1995*. Madrid: Ediciones Tempo.

RAE (s. f.) 'Divulgar'. Disponible en: https://dle.rae.es/divulgar (Consulta: 11 enero 2024).

RAE (s. f.) 'Vulgarizar'. Disponible en: https://dle.rae.es/vulgarizar (Consulta: 11 enero 2024).

RAE (s. f.) 'Realidad'. Disponible en: https://dle.rae.es/realidad (Consulta: 5 enero 2025).

Rogers, E. M. y Dearing, J. W. (1988) 'Agenda-setting research: Where has it been? Where is it going?', *Communication Yearbook*, 11, pp. 555–594.

The Guardian (2017) 'C. P. Scott centenary essay'. Disponible en: https://www.theguardian.com/sustainability/cp-scott-centenary-essay (Consulta: 7 enero 2025).

Venegas Ramos, A. (2023) *Pantallas de la memoria. Cómo y por qué las imágenes digitales transforman nuestra idea de la historia*. Madrid: Clave Intelectual.

CUADROS
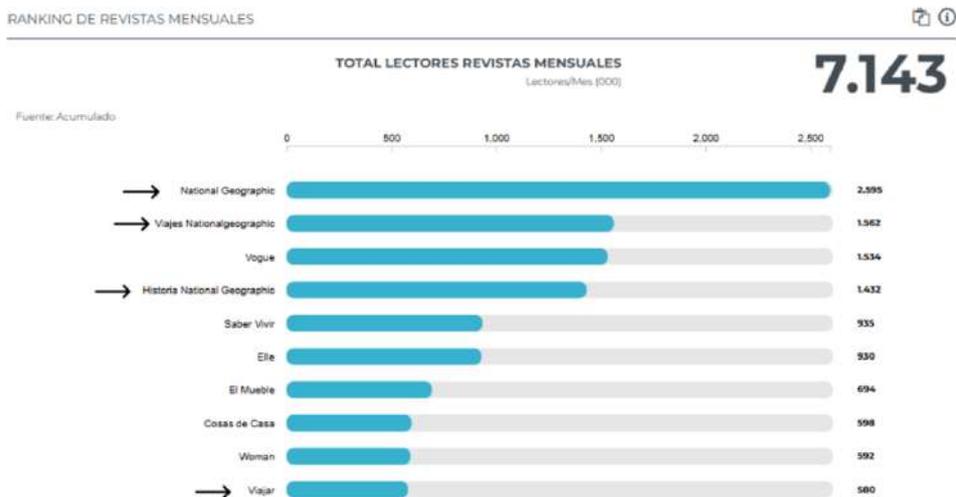
Cuadro 1. Revistas temáticas mensuales más leídas según el EGM 2025, con presencia destacada de cabeceras de contenido histórico o patrimonial como indicador de interés social por la disciplina.

Cuadro 2. Esquema de propagación del conocimiento histórico desde la academia hasta el público general, mostrando las mediaciones que pueden generar simplificación, distorsión o vulgarización.

Divulgar en femenino. Desde la historia de las mujeres a ser divulgadora en redes

Knowledge in the Feminine: From Women's History to Digital Public Engagement

Patricia González Gutiérrez
Desperta Ferro Ediciones
patriciagonzalezgutierrez@gmail.com
ⓘ 0000-0002-9342-8609

La historia de las mujeres y de género ha tenido que abrirse paso en un mundo muy masculinizado. Se suele decir que el inicio de la misma está en la escuela de los Annales, aunque no es realmente cierto. Ni la historia de las mujeres ni las historiadoras de dicha escuela tuvieron realmente un apoyo entusiasta por parte de sus pares. Hubo que esperar al empuje feminista para que, realmente, se empezara a hablar de una historia en femenino, de una historia de la diversidad, de una historia de colectivos tradicionalmente feminizados, como el LGTBIQ (González Gutiérrez, 2025).

Precisamente el influjo de los feminismos en el desarrollo y consolidación de la historia de las mujeres y de género ha provocado también reacciones machistas y una acusación de politización. Las nuevas perspectivas, que en su momento se despreciaron como un interés en cuestiones banales o menores, se empezaron a ver como ataques a una forma de hacer historia más tradicional. Y, sobre todo, cuando estas abandonaron la torre de marfil y permearon a la divulgación, de una forma u otra, se vieron como un ataque al orden social.

1. Una historia en masculino no genérico

La historia positivista era una historia masculinizada. Las mujeres han estado prácticamente ausentes, salvo por excepciones, usadas como modelos o contramodelos. Uno de los primeros libros sobre romanas, el de Roergas de Serviez, sobre las mujeres de la casa imperial, *Les Impèratrices Romaines*, de 1720, se explaya en el prólogo con una disculpa por lo irrelevante del tema y se justifica diciendo que lo escribe para prevenir a los hombres de las maldades femeninas (Cid López, 2010, pp. 685–702).

Lo mismo se aplicaba a la divulgación y la representación histórica en las instituciones. En muchos museos, en que se conservan proyectos de musealización

antiguos o "tradicionales" podemos percibir aún esta exclusión de las mujeres (la representación gráfica de los árboles o líneas evolutivas es, quizás, una de las imágenes más representativas), infancias, colectivo LGTBIQ, la discapacidad o racialización; también de los elementos problemáticos, tanto de los que destacan los elementos más "sucios" del mundo bélico como de los elementos que se salen de nuestros valores, como los genocidios, la pedofilia e incluso la esclavitud. Se crearon, así, museos "aptos para todos los públicos" que escondían los problemas sociales y crean un relato aséptico, o incluso propagandístico, de las conquistas y movimientos de pueblos.

Por otro lado, la mayoría de estas imágenes se crearon, como tópico, en el siglo XIX y trasladaban la idea de la sociedad decimonónica al pasado. Un buen ejemplo son las imágenes de mujeres en cuevas, no haciendo nada, criando o recolectando, mientras los hombres aparecen cazando, pintando o tallando. La imagen es tan potente que se ha trasladado a todas las imágenes de la prehistoria, incluso las de los juguetes, como podemos ver, por ejemplo, en la imagen elegida para promocionar el set dedicado a la prehistoria de Playmobil.

Las nuevas corrientes historiográficas han supuesto una renovación no solo en la historia escrita, sino también en su transmisión al público, y de ello es un buen ejemplo la transformación paulatina de los proyectos museísticos (Hornos and Rísquez, 2000, pp. 175–186; Querol, 2013, pp. 63–80). Cuestiones como la ausencia de las mujeres en las representaciones gráficas, el olvido de la infancia o la discapacidad o la exhibición de restos humanos o la "folklorización" de otras culturas mediante un discurso colonial se han puesto en los últimos años sobre el tapete, no sin polémica (Gifford-Gonzalez, 1993, pp. 3–21; Fabien van Geert, 2020).

Todo ello ha tenido que dar luego el salto a la divulgación generalista como tal, con proyectos en forma de charlas, libros, podcast, cuentas en redes sociales... que se enfrenta, aún más que en la academia, a una serie de prejuicios profundamente enraizados en la sociedad. La divulgación española, durante mucho tiempo, fue algo que se hacía, en general y salvo honrosas excepciones, como la revista *Historia 16*, de forma independiente de la academia, por parte de personas que no eran historiadoras ni tenían formación específica, como un *hobby* o dentro de programas generalistas. La ANECA, encargada de evaluar becas y acreditaciones, durante mucho tiempo fue reacia a reconocer la divulgación como parte del trabajo académico, por lo que alejó, al contrario de lo que sucedía, por ejemplo, en los países anglosajones, a los investigadores de la diseminación del conocimiento a un público amplio.

Todo ello está cambiando a bastante velocidad en los últimos años. Los documentales eran una de las grandes excepciones a esta dinámica, aunque funcionaban de una forma autónoma y particular, sin continuidad en el tiempo. De hecho, es el principal medio en muchas regiones y entre muchos grupos sociales

para el conocimiento histórico (O'Connor, 2001, pp. 1200–1209; Paz, 2015, pp. 275–302; Hernández Corchete, 2008). A este medio audiovisual se ha unido el uso de las nuevas tecnologías en la difusión del patrimonio o los podcasts históricos. Asimismo la recreación histórica, que se ha ido organizando en Europa desde los años ochenta y que llegó a España con algo de retraso, ha ido evolucionando de un ámbito eminentemente militar y con cierta dosis de banalización de la historia a un medio efectivo de divulgación, con una mayor recreación de la vida cotidiana, del contexto social y con un asesoramiento cada vez más profesional (Español y Franco, 2021).

El machismo, además, vende. La perpetuación de mitos misóginos como los de la supuesta ninfomanía de las mujeres romanas, egipcias o griegas, o las historias de gobernantes descontroladas sexualmente, siguen siendo un *bait* en medios escritos y visuales. Desde las películas sobre Mesalina como una femme fatale hasta la idea de la "perversión romana" que se refleja en el imaginario colectivo o la obra de autores como Carcopino (Carcopino, 2001), los mitos históricos misóginos han sido una costumbre poco cuestionada. Así, en 2018, podíamos leer en *El Español* un titular que rezaba "Cleopatra, la experta en felaciones conocida como 'la boca de 10.000 hombres'". El artículo no mejoraba el titular, sin que, por supuesto, se aportara ningún documento que respaldara las peregrinas afirmaciones de quien lo escribió (*El Español*, 2018).

2. IDEOLOGÍAS VISIBLES E INVISIBLES

Como hemos dicho, las propuestas de renovación y aplicación de nuevas perspectivas no han sido siempre bien recibidas, y ciertos sectores de la población han expresado burlas y quejas por todo lo que se saliese del discurso positivista y más tradicional. "Politización sectaria" o "basura woke" son términos repetidos por quienes quieren evitar debates sobre cómo se construyen los discursos históricos. De nuevo, la acusación sobre el sesgo ideológico, resumido como "la ideología", obvia los problemas que ha supuesto una visión colonial, racista, androcéntrica o LGTBIfóbica en la representación de nuestro pasado, invisibilizando aquellas ideologías que han sostenido el orden social hegemónico en cada época. Al final, cualquier reconstrucción de nuestro pasado supone tomar una serie de elecciones, sobre todo cuando se trata de qué se expone o qué se oculta (Aparicio, 2025).

También ha habido problemas con algunos intentos de contextualizar a ciertos autores o las circunstancias de ciertas obras, sobre todo en las pinacotecas, tradicionalmente más conservadoras en su discurso museográfico que los museos arqueológicos o antropológicos. Aun así, han surgido debates en torno a artistas especialmente problemáticos, como el caso de Gauguin, en torno a cómo exponer obras en que entran temas tan delicados como la pederastia y el abuso sexual (Elephant, 2020; Independent, 2019). Mucho se ha enfatizado en los últimos años, por ejemplo, cómo la forma de exponer y nombrar las escenas de abuso sexual y violación en el arte contribuye a la normalización de dichas violencias. Podemos

pasar sin el más mínimo sentimiento de incomodidad, solo apreciando la belleza de la obra, por salas que muestran una violación tras otra.

El imaginario colectivo tiene un claro sesgo masculino, algo que tiene causa y consecuencia en que la historia divulgada haya sido tradicionalmente positivista, con un foco muy fuerte en figuras consideradas heroicas, batallas y movimientos políticos. Las figuras de las mujeres han sido escasas, estereotipadas e incluidas en esas "mujeres fuertes" que sirven, además, de descarga de responsabilidad (¿como va a ser la historia androcéntrica si hablamos de Isabel la Católica y María Pita?). Esto no viene de la nada, y ha existido un círculo vicioso entre cómo se ha estudiado la historia desde la Academia, cómo se ha utilizado por parte del Estado, partidos y políticos y cómo se ha trasladado al público a través del arte, los audiovisuales o la divulgación.

Aunque esta visión no era inocente en muchos casos, naturalizaba unas exigencias especiales si se quería hablar de mujeres u otros colectivos minorizados. En abril de 1985, Lawrence Stone reseñó dos libros relacionados con la historia de las mujeres para *The New York Review of Books. Se* sintió en la profunda obligación de, antes de pasar a reseñar el trabajo de dos compañeras de profesión, realizar un decálogo explicándoles cómo hacer su trabajo (Stone, 1985). El primer mandamiento era no escribir sobre mujeres si no se relacionaba también con los hombres o las infancias, porque es necesario conocer las interacciones entre ellos. También decía que no se debía confundir la norma con la realidad social o que no se debían usar tres o cuatro biografías para describir la experiencia de la mayoría. Asimismo, decía, no se debía usar la historia para apoyar el feminismo y recordar que los cambios no los realiza una minoría y que son siempre relativos. Perfecto, diríamos... pero ¿Se aplicaba el ilustre autor sus propias normas? ¿Por qué especificaba que eran normas aplicables a la historia de las mujeres? ¿Alguien jamás las explicitó o aplicó a la historia "masculina"?

Las respuestas parecen evidentes. La historia se había construido tradicionalmente sin mujeres ni niños y a nadie le había importado. La historia había sido tradicionalmente la de los grandes hombres, con una idea de que eran los reyes y generales quienes cambiaban el mundo. Las fuentes se habían leído al pie de la letra sin plantearse que la realidad que nos muestra la arqueología no siempre coincidía con lo que nos decían los grandes moralistas. Y aún sigue pasando. La Historia, con mayúsculas, era la de los hombres y la política, con la que todo el mundo debía sentirse identificado (e identificada), mientras que el resto eran historias particulares y diminutas. Todo el mundo debía considerar a César como su pasado, pero solo las mujeres podían ver en Livia a su antecesora. Y este ha sido el punto clave que ha permitido obviar, ignorar o despreciar la investigación y la divulgación "en femenino".

Ha existido, además, tradicionalmente (y aún se conserva), un doble rasero en que las cuestiones relacionadas con la historia de las mujeres deben probarse hasta la extenuación, sin posibilidad para la teoría, y con una certeza más allá de lo razonable, mientras que nadie jamás pidió dichas pruebas para lo que se consideraba "de sentido común" (González Gutiérrez, 2018, pp. 55–67; Sánchez Romero, 2022). Así, mientras que hace falta que una cazadora prehistórica se aparezca en un museo para explicar cómo abatía ciervos, nadie jamás pidió pruebas para afirmar que los hombres cazaran, o pintaran o realizaran industria lítica. Lo mismo pasa con el trabajo femenino, sobre todo si se asocia a lo físico o el prestigio. Necesitamos muchas más pruebas que para cualquier afirmación relacionada con la historia de los grupos "normativos". Parte de estas estrategias son intencionadas, en lo que en el mundo anglosajón ha venido llamando *sealioning*, es decir, una estrategia para hastiar al interlocutor basada en las preguntas malintencionadas, el cambio de foco cada vez que se agota una rama, la repetición de preguntas, el rechazo de las respuestas y la petición incansable de fuentes y pruebas. De hecho, el nombre viene de un cómic que ejemplificaba con un león marino parlante la estrategia (Wondermark, s. f.).

Las redes son un ejemplo perfecto de la proliferación de las estrategias malintencionadas. El anonimato y la capacidad de interacción automática, sin reflexión, permiten ver algunos de los sesgos más asentados, pero no siempre visibles. La negativa, por ejemplo, a aceptar que las mujeres cazaban en la prehistoria o trabajaban en el mundo clásico viene, precisamente, de un prejuicio ideológico y, por tanto, se niega basándose en supuestos ejemplos antropológicos (cuando, en realidad, en muchos grupos modernos las mujeres sí cazan), o en una supuesta tradición de domesticidad generalizada que solo se ha dado en ciertas épocas y grupos sociales privilegiados, mientras que la mayoría de mujeres han tenido que trabajar en los talleres familiares o en trabajos externos (Anderson *et al*., 2024; Alfaro Giner, 2010, pp. 15–38; González Gutiérrez, 2021, cap. 6; Segura Graiño, 2001, pp. 109–120). Los datos se vuelven irrelevantes en este tipo de debates o, en muchas ocasiones, se niegan o ignoran.

3. LA HISTORIA ES "COMO LA PAELLA"

Todos estos sesgos androcéntricos y las reacciones violentas a los cambios de paradigma en la historiografía no nacen del vacío. La asociación entre pasado y presente hace que el análisis de las dinámicas sociales del pasado que hoy (o incluso en ese momento) se considera(ba)n perversas, se tome como un ataque personal, sobre todo aquellas que han dejado un claro rastro en el presente. ¿Por qué el título de este apartado? Porque es comparable a esas cuestiones que son claramente identitarias, pero pasan desapercibidas en la vida cotidiana, como la forma de comer. Las reacciones humorísticas o indignadas ante un inglés que echa chorizo o piña a la paella dejan de ser inocentes cuando nos planteamos, por ejemplo, por qué se responde sistemáticamente en redes o en grupos sociales al veganismo con fotos

de asados, salchichas o filetes. La agresividad, en esos casos, sea ante una persona que realiza un activismo o simplemente frente a un particular que comparte una receta, nos dice mucho de cómo se percibe la identidad y, en este caso concreto, la masculinidad. Precisamente la vinculación tradicional de la carne con una masculinidad normativa hace que posibilidades como el veganismo sean vistas no como opciones, sino como ataques (Adams, 1990).

Todo ello en un ambiente en que la recreación de los pasados violentos tiene un público especialmente amplio, a través no solo de las series, videojuegos o películas, sino también de la propia forma de musealizar los restos materiales de las guerras, conquistas y represiones o del propio paisaje urbano y sus lugares de memoria, como las esculturas (Grever y Van Nieuwenhuyse, 2020, pp. 483–500; Venegas, 2023).

En ocasiones se llega a negar la existencia o la capacidad de formar parte de la historia de ciertos grupos. La indignación ante una serie sobre Alejandro Magno en que se afirmaba que se había "hecho gay" a Alejandro es un buen ejemplo. También algunos comentarios y comparaciones en redes sociales. Uno de ellos rezaba "Y de los extraterrestres quien se acuerda e? (sic) Quién? Otro colectivo discriminado por los arqueólogos" en respuesta a un tuit de *El País* que enlazaba una entrevista a Marga Sánchez Romero en que reflexionaba sobre la necesidad de estudiar a las mujeres en la historia. De nuevo, el retorno a calificar de "woke" o invención a todo aquello que se salga de la historia más tradicional y positivista.

Es paradójico que la divulgación de la historia de las mujeres o de otros colectivos sea, además, calificada de presentismo. El término se usa sin demasiado conocimiento sobre el significado real del mismo. De hecho, se suele aplicar para evitar una crítica a cuestiones del pasado que hoy resultan incómodas a ciertos grupos, como la legalidad de ciertos tipos de violación o violencia de género, la esclavitud o el exterminio de pueblos. Esa negación de la alteridad moral del pasado, que lleva a una identificación personal con el mismo resulta peligrosa, y conlleva la invisibilización de aquellos aspectos que no gustan o que, incluso, parecen estéticamente feos. No hay mayor ejemplo de presentismo, al querer igualar los valores actuales con los de sociedades pasadas. Para la negación del análisis histórico también se tiende a negar la capacidad de aplicar categorías de análisis modernas, como el género al pasado porque "no se tenía el concepto", ignorando dicho reclamo en otras cuestiones como la idea de nacionalidad o Estado, o periodización (cuando se habla de Reconquista o Edad Media, por ejemplo)

El mundo clásico ha sido tomado siempre como un *exemplum*, como un modelo idealizado, en Occidente, pero las nuevas producciones remarcan ese imaginario, como pasa en el caso de la película de 300, de Snyder. Hay que entender que la fórmula *historia magistra vitae* no es vista como una forma de entender el mundo y de dónde vienen los usos sociales presentes (como se concibe la historia como

disciplina actualmente), sino como una inspiración, que tiene que ser épico y luminoso (Hodkinson, 2022, pp. 59–84). Como con las fuentes antiguas, la importancia real de la veracidad y el rigor se diluyen en favor de un relato épico, de unas consignas políticas o de un sentimiento de comunidad. En este marco, las historias más "sucias" o que estudian la violencia o los grupos discriminados se ven, como hemos dicho, como un ataque a la identidad y el orden social.

4. LA EXTREMA DERECHA Y LA APROPIACIÓN DE LOS CLÁSICOS

En la actualidad, además, hay una clara fagocitación del mundo clásico por parte de la extrema derecha y los movimientos identitarios (du Bois, 2021). De un mundo clásico, por supuesto, cargado de épica y testosterona. Gladiadores y soldados, legionarios y emperadores, espartanos y romanos... En parte la simbología tiene asociaciones claras. El uso, por ejemplo, del lema "molon labe" y la lambda, así como la batalla de las Termópilas en general, se asocian a un supuesto momento de "freno" a los persas, asociados a lo oriental, inmigración como conquista (Naerebout, 2022, pp. 59–84).

En Estados Unidos la apropiación y uso de la imagen de militarismo espartano y romano ha sido especialmente visible. En el asalto al capitolio de enero de 2021 varios de los asaltantes llevaban cascos y motivos inspirados en la antigua Grecia, y la congresista republicana Marjorie Taylor Greene llevó en algunas ocasiones una mascarilla con el lema de "molon labe". Otras imágenes que nos dejó el asalto, además de los movimientos posteriores, fueron los lemas, pancartas y memes que asociaban a Trump a César cruzando el Rubicón, que se hicieron con IA o superponiendo su cara a imágenes de Máximo Décimo Meridio, de Gladiator.

El grupo neofascista *Identity Evropa* también recurrió al arte clásico para ilustrar sus carteles propagandísticos, bien con esculturas del mundo antiguo o de personajes del mismo. Tampoco en España nos hemos librado de ese tipo de imágenes y se hizo viral una imagen que asociaba a los líderes de VOX con los espartanos, usando una imagen de la película *300* (El Periódico, 2019). También inició su campaña de 2019 en Covadonga, aludiendo a un mito fundacional, o colgó una foto en su cuenta oficial de Santiago Abascal con un morrión, en respuesta a un tuit de Pablo Casado con la llamada "cruz de la victoria" de Oviedo en que se alzaba como defensor e iniciador de una nueva "Reconquista". En 2016 VOX ya había sacado un video promocional en que una de las escenas recreaba la famosa imagen de Gladiator en que el protagonista pasa la mano por un campo de trigo.

Aquí entra también la influencia del cine en particular y los medios audiovisuales en general. El nuevo furor por los espartanos deriva mucho más, precisamente, de la película de *300*, de Zack Snyder, que de una recuperación de la tradición clásica del "mirage espartano". Lo mismo pasa con las nuevas producciones bélicas, en un sentido amplio, centradas en Roma, ya sea *Gladiator*, las series de *Spartacus* o *Roma* (Aguado, 2020).

Si la historia se usa como elemento identitario y unificador en torno a una ideología, el machismo no resulta tampoco una gran ausencia. Una de las imágenes más populares de la campaña presidencial de Donald Trump en 2016 fue la que le representaba como Perseo con la cabeza de Medusa, caracterizada como Hillary Clinton, en una imagen que hace aflorar no solo componentes políticos de apropiación de un pasado clásico, sino unas grandes dosis de misoginia y violencia contra las mujeres. En un contexto de exaltación del militarismo, la presencia femenina resulta complicada, aunque sea necesario incluirlas, y el androcentrismo se vuelve un elemento central que, en muchas ocasiones, resulta en agresividad.

No solo en las imágenes vive este sesgo androcéntrico. "Phalanx" tienda de ropa y *merchandising* de movimiento extrema derecha, con símbolos como cascos griegos unidos a "movimiento identitario". De hecho, en algunas de estas tiendas se nota el enfoque claramente masculino. Si comparamos en algunas de ellas, como puede ser *Ansgar Aryan*, la cantidad de productos, como la ropa, enfocados a hombres y mujeres, los primeros son abrumadoramente más en número y variedad.

En esta y otras tiendas de este estilo puede observarse que la cantidad y variedad de ropa y objetos dirigidos a las mujeres es mucho menor, por lo general, que los enfocados a un público masculino. El público objetivo está claro, pero también es un elemento consciente e ideológico.

Cuando la divulgación y la investigación sobre el mundo clásico "desmontan" o se alejan de esa imagen guerrera y masculina, cuando se habla de mujeres romanas, de discapacidad, de infancias, de vida cotidiana, de tragedias y se expone el lado más oscuro de la violencia, no se entiende como un complemento o como un estudio válido, sino como una agresión a esa identidad.

En España la apropiación de los clásicos compite (y no siempre gana) con los mitos y elementos históricos identitarios medievales y modernos, como Covadonga, la Reconquista o el imaginario en torno a los tercios, como hemos visto con el ejemplo de las campañas de VOX. Por un lado, son mitos nacionales que han sido ampliamente exaltados por la dictadura, mientras que el pasado romano es ambiguo, con la frecuente exaltación de los episodios de resistencia como Numancia. Por otro lado, la carga clasista en torno a lo que en el mundo anglosajón define como *classics* (una mezcla de filología clásica e historia antigua) es mucho menor en nuestro país (Skinner, 1987, pp. 181–186). Hay que tener en cuenta que John C. Calhoun, en el siglo xix, había llegado a decir que solo el día que conociera a un negro que supiera de sintaxis del griego clásico creería que los negros eran humanos. Asimismo, existe una complicada relación entre la historia y la investigación de la misma. Para la reapropiación se hace necesaria la exaltación del imaginario colectivo y la vocación no profesional. Por otro lado, requiere descartar la investigación profesional, por su complejidad y diversidad de aproximaciones y por su labor de deconstrucción de los mitos.

5. DIVULGAR DESDE LOS MÁRGENES

Divulgar historia de las mujeres, especialmente dentro de la historia antigua, como hemos visto, supone enfrentarse a muchos problemas. Sin embargo, hay un segundo factor que marca decisivamente la divulgación, y es el género de quien divulga. Ser mujer, en estos casos, supone un factor diferencial, como también lo es pertenecer a grupos que se consideran feminizados, como el colectivo LGTBIQ.

Aquí surgen problemas que son comunes a otras ramas de la divulgación científica. De hecho, cuando se pregunta a distintas divulgadoras, las respuestas suelen ser parecidas. Un ejemplo es el video que hizo Emily Graslie, que llevaba en ese momento un canal de divulgación para el *Field Museum of Natural History* de Chicago, respondiendo a la pregunta de si había sufrido sexismo en su labor de divulgadora científica (enlace disponible en *Youtube*). Habló sobre la escasez de mujeres en ciertos campos de divulgación, sobre cómo era un problema similar al de los techos de cristal en otros ámbitos laborales o sobre los comentarios desagradables o inquietantes que recibía.

También han surgido, con una fuerte base *online*, movimientos de exaltación de un tipo de masculinidad, en cierto modo, relacionada con la masculinidad hegemónica tradicional, mezclada con componentes de misoginia exaltada, resentimiento y con la cosificación de las mujeres y el culto al cuerpo. Estos han recurrido, en los últimos años, a una identificación con una cultura clásica hipersimplificada, sobre todo con el estoicismo. Esta se reduce, muchas veces, a frases motivacionales y a un concepto de individualidad, más que al espíritu de la escuela filosófica. Aunque Marco Aurelio se lleva la palma entre las citas de estos grupos, también se recomiendan obras de Séneca o Epicteto (Zuckerberg, 2019, pp. 45 ss.). Algunas de estas comunidades misóginas, como la incel y otros grupos de la "manosfera", han hecho del acoso en redes y el ataque a las mujeres un elemento fundamental de su identidad (Mantilla, 2015). Campañas de acoso como la conocida Gamergate, de 2014, son cada vez más frecuentes. La exposición que requiere la divulgación o el activismo en redes pone a las mujeres en su punto de mira.

Asimismo, parte del ataque a lo "woke" es identificar la educación básica o el respeto a la dignidad y los derechos de la gente como "corrección política" e intentar convertir el insulto en una especie de seña de identidad. Entra dentro de lo que Mauro Entrialgo definiría como "malismo" (Entrialgo, 2024). Así pues, el uso del acoso on line, el insulto y el "trolleo" se han convertido en una herramienta más para ciertos grupos, que proporciona a sus individuos validación de grupo. Y, en muchos casos, permite monetizar estas actitudes, a través de las redes sociales o al crear una comunidad entregada dispuesta a tragar con lo que sea, sean mitos históricos o actitudes de acosador de colegio.

Los insultos sobre el físico o la (supuesta) vida sexual, las amenazas o deseos de violación o asesinato, o los acosos continuados en el tiempo son características de lo

que se ha venido a llamar "troleo de género" o "gendertrolling". Las mujeres con una cierta exposición en redes, sea por el motivo que sea, sufren este tipo de acoso de forma habitual, y este aumenta cuando se combina con ciertas características de su exposición, como la divulgación en temas "sensibles" para ciertos grupos o sobre la existencia misma de dichos grupos. Así pues, la actividad en redes sociales, que permite interactuar directamente con las investigadoras, se presenta como un lugar especialmente potente de divulgación, con la posibilidad de compartir la información, resolver dudas y ampliar la información, pero también un lugar especialmente sensible a la violencia (Piñeiro Otero *et al.*, 2024).

En todo ello juega un papel importante, también la impunidad que proporcionan las redes sociales, no solo por el anonimato que permiten, frente a las víctimas, que suelen mostrar su nombre y apellidos, sino también por los escasos medios de denuncia y control de la violencia. En algunas, como el actual X, antes Twitter, los bloqueos han perdido eficacia, al no impedir la visibilidad para la persona bloqueada, y resulta enormemente complicado que se penalicen los tuits machistas, lgtbifóbicos o racistas (EXIST Project, 2021). Esto permite que cuentas dedicadas en exclusiva o de forma activa al odio y el acoso campen a sus anchas sin mayores consecuencias, pese a las normas que, en teoría, penalizan dicha actividad. De hecho, fue el caso del acoso y las amenazas a Caroline Criado-Pérez, que recibió a miles tras pedir al gobierno británico que incluyera más imágenes de mujeres en los billetes, lo que llevó a algunas redes sociales a facilitar un medio de denuncia en las propias plataformas, con un éxito relativo y la advertencia sobre que no serviría para "simples insultos". Se calcula que más del 70% de las víctimas de abusos *online* son mujeres y que por cada cien mensajes sexualmente explícitos que reciben las mujeres, los hombres recibirían menos de cuatro (PSMag, 2013).

La libertad que proporciona un espacio que permite que solo importen las palabras y que hizo a Donna Haraway escribir "prefiero ser un ciborg que una diosa" (Haraway, 1984, pp. 65–107), tiene su cara más oscura en la libertad que proporciona para ejercer violencia sin consecuencias. "Gorda", "Charo", "No sirves ni para ser violada", "vaca", "Escoria"… Son insultos que pueden verse a diario, y que caen, una y otra vez, sobre las diversas divulgadoras. Los insultos físicos a los hombres, aunque pueden afectar al físico ("calvo", "gordo", entre otros) normalmente se centran más en la feminización (y quizás no debamos obviar cómo la estatura, la calvicie o el sobrepeso se usan como elementos de feminización), con adjetivos como "sojas" o "planchabragas". También son habituales las amenazas directas, bien en forma de amenazas y fotos sexuales o comentarios como "te vamos a matar", "a ver si te dejan ir a una excavación pero ten cuidado si no topas con una mina antes (sic)". Asimismo, la modificación de fotos de las divulgadoras como forma de burla o amenaza también se ha convertido en un recurso habitual.

Las campañas de acoso pueden prolongarse en el tiempo. Es decir, no son solo acciones puntuales de particulares, sino que pueden durar semanas, meses o años

(como pudo verse en el Gamergate), y desarrollarse de forma coordinada. Hay personas que pueden seguir durante meses a una divulgadora insultándola de forma habitual, cambiando de cuenta si se la llegan a bloquear o eliminar, o personajes con gran capacidad social que pueden dirigir a sus seguidores para que violenten a ciertas personas en concreto. Helen Lewis, una de las primeras periodistas en hablar abiertamente de la misoginia online acabó por dar nombre a la llamada "Ley de Lewis", que afirma que todo post sobre feminismo se justifica por los comentarios al mismo. La cantidad de comentarios insultantes o agresivos se ve también en cualquier artículo, post en redes o reseñas de libros que hable de historia de las mujeres o de género.

En un rango menor pero que, en casos extremos, ha tenido consecuencias poco deseables, está lo que se conoce como "relaciones parasociales", en que los seguidores de un personaje conocido consideran que existe un vínculo emocional que, en realidad, no se da. Esto puede llevar a la exigencia de atención, a las amenazas en caso de no cumplirse las expectativas o la decepción ante una respuesta indiferente.

Mary Beard ha sido especialmente visible en su denuncia ante estas actitudes y esta forma de trolleo. También sobre las estrategias de silenciamiento de las mujeres en otros campos y cómo se aplican dentro y fuera de las redes sociales (The Guardian, 2014). Su posición como académica, unida a su exposición en medios audiovisuales y su negativa a presentarse de una forma "normativa" han sido una combinación explosiva. Su manifiesto *Women & Power* se inicia con un capítulo titulado "*Public voice of women*" y una historia muy significativa, la de Telémaco haciendo callar a su madre, Penélope, y mandándola a sus habitaciones (Beard, 2017). Es una historia que, sin la relación familiar de por medio, se ha repetido una y otra vez en la historia de la humanidad. Como bien dice, la mujer que se atreve a hablar en público, incluso si no es silenciada, tiene que pagar un alto precio.

Por otro lado, en un mundo tradicionalmente masculinizado como es el de la divulgación (y ha sido la Academia hasta hace muy poco), resulta habitual que se den los conocidos como "*all male panel*", es decir, congresos y eventos en que no hay ninguna mujer... incluso en aquellos en que se habla de temas relacionados con las mujeres. Asimismo, hay una cierta tendencia a la "tokenización" de las investigadoras y divulgadoras, precisamente para evitar los reproches a ese tipo de situaciones, y a crear eventos, conferencias y charlas en que se invita a mujeres exclusivamente en fechas destacadas como el 8 de marzo o el 25 de noviembre, o a hablar de "ser mujer en".

Las mujeres, además, pierden el "nombre" en muchas ocasiones. Es algo que ya se ha estudiado en torno a la comunicación periodística, cómo ellos son citados por su nombre y apellidos mientras ellas se convierten en "una mujer" o "la mujer de" o, peor, en "una chica". También existe la tendencia a pensar que todo artículo o libro

firmado por un pseudónimo o en que solo se cita la inicial del nombre es de un hombre. Como diría Virginia Woolf, anónimo siempre ha sido nombre de mujer. De nuevo pasaré a la primera persona. Tras una polémica con un conocido comunicador sobre si en la Armada británica había o no personas negras en época napoleónica, a raíz de una escena en una película, este hizo un video. En el mismo se citaba a las dos personas con quien más se había enganchado. Mientras al otro divulgador le citó por el nombre, yo me vi reducida a "una historiadora".

La condescendencia de ciertos hombres al hablar con mujeres cristaliza en lo que se ha llamado "mansplaining" o "machoexplicación". Es decir, hombres que, haciendo gala del efecto Dunning-Kruger, intentan explicar a mujeres el propio campo en el que ellas son expertas. De hecho, el término nace del libro de Rebecca Solnit *Men Explain Things to Me* (Los hombres me explican cosas), en que narra como un hombre le intentó explicar un libro del que ella era autor y que él, obviamente, no se había leído. Aquí puedo pasar a la primera persona y contar algunos de los comentarios más hilarantes recibidos en redes (en un sentido amplio). Un señor, en un comentario a una entrevista sobre el libro *Soror*, me intentó explicar que no tenía ni idea de lo que hablaba porque ¿cómo iba Livia a haber sido madre y abuela de emperadores si no tuvo descendencia con Augusto y "vivió 29 años" (sic)? Tardé un rato en darme cuenta que no había siquiera manejado bien las cifras antes y después de Cristo y por eso le salía esa edad. Aun así, intentaba darme lecciones. Otro señor intentó darme lecciones sobre por qué Cleopatra sí era una "puta" porque se había acostado con "el tío y luego el sobrino". Tardé otro rato en darme cuenta de que se refería a César y Marco Antonio, de los que había confundido a sus ascendientes femeninas. Otro dijo que no tenía ni idea sobre prehistoria porque ¿cómo iban a cazar las mujeres si no había "bifaces adaptados" (sic)? Se enfadó mucho cuando le puse la foto de una vitrina con bifaces de distintos tamaños. Más de una divulgadora ha sufrido lo mismo que Solnit, sobre todo cuando escribía blogs en que no quedaba clara la autoría o cuando se firma solo con la inicial. De hecho, la reacción ante un contenido varía cuando la autora es visible frente a una escritura más anónima. Aquí cabe recordar a Laia San José Beltrán, que lo ha sufrido frecuentemente en relación a su trabajo en el blog de divulgación *The Valkyrie's Vigil*, sobre vikingos, también un tema muy masculinizado.

El "sentido común", en estos casos, suele jugar malas pasadas y, en muchos casos, se vuelve una estrategia consciente. Si siempre ha sido así, si cuela, si existe el mito, lo que puedas pensar se convierte, automáticamente, en la realidad, sin que sea necesario respaldar la afirmación con ningún dato, siquiera, informarse. El sesgo de confirmación de "¿cómo no va a ser verdad si coincide con lo que quiero?" es peligroso y crea historias brillantes, épicas y cómodas... aunque sean falsas. Esto, en parte, además de la cuestión identitaria, explica la dificultad para desmentir mitos históricos recurrentes en el imaginario colectivo, como que Isabel la Católica abolió la esclavitud, el PSOE votó en contra del sufragio femenino o que las mujeres no trabajaban hasta el siglo XX.

6. CONCLUSIONES

Un mundo de noticias falsas necesita poner en duda a los expertos, un mundo de discriminación necesita callar a los discriminados, un mundo en que ciertos sectores pretenden apropiarse la historia, necesita de violencia contra quien se sale de esas líneas.

La combinación entre el uso político de periodos históricos como el mundo clásico y la necesidad de borrar ciertos aspectos del mismo para convertirlo en algo útil, el auge de los movimientos misóginos en redes y la tradición masculinizada de la Academia y la divulgación complican una divulgación en femenino. No solo cuesta más llegar al gran público, sino que requiere que las divulgadoras paguen un precio en cuanto a salud mental e incluso física.

No todo es terrible. Las cosas han cambiado mucho y también existen grandes comunidades, nuevos espacios de divulgación y redes de apoyo. En la Academia hace décadas que la historia de las mujeres y de género no se cuestionan (salvo excepciones) y que ha aumentado la consideración de la divulgación y su necesidad. En la divulgación cada vez más se es consciente de la necesidad de la diversidad de puntos de vista y de la presencia femenina. La recreación cada vez más abarca la vida civil y de las mujeres y ya no se limita a lo militar. Cada vez más público accede a una divulgación histórica de calidad y diversa. Cada vez hay más divulgación sobre la historia de las mujeres y más divulgadoras.

Pero no podemos cerrar los ojos ante violencias y problemas que aún siguen presentes, precisamente porque es nuestro campo de estudio. La frase de que toda historia es historia contemporánea nunca tuvo tantos matices. Necesitamos buscar redes de socialización y apoyo, soluciones a las campañas de acoso y entender cómo funcionan estas dinámicas para que las historiadoras, investigadoras y divulgadoras no sigan pagando un alto precio por su trabajo y su exposición.

BIBLIOGRAFÍA

Adams, C. J. (1990) *The sexual politics of meat: A feminist-vegetarian critical theory*. New York: Continuum.

Aguado, O. (2020) 'El "efecto Gladiator" 20 años después: cine "de romanos", Champions, memes y extrema derecha', *Proyecto ANIHO – ANIWEH Project*. doi: 10.58079/b5uk.

Alfaro Giner, C. (2010) 'La mujer y el trabajo en la Hispania prerromana y romana', *Mélanges de la Casa de Velázquez*, 40(2), pp. 15–38.

Anderson, A. et al. (2024) 'Correction: The myth of man the hunter: Women's contribution to the hunt across ethnographic contexts', *PLOS ONE*, 19(8), e0309543.

Aparicio, P. (2025) *Ventanas al pasado. La reconstrucción de la historia con imágenes*. Madrid: Akal.

Beard, M. (2017) *Women & power. A manifesto*. New York: Liveright.

Carcopino, J. (2001) *La vida romana en el apogeo del Imperio*. Madrid: Temas de Hoy.

Cid López, R. M.ª (2010) 'Mujeres "poderosas" del Imperio romano en la historiografía moderna. Algunas notas críticas a las visiones de la Ilustración y su influencia', in Fornis,

C. et al. (eds.) *Dialéctica histórica y compromiso social: homenaje a Domingo Plácido*, pp. 685–702. Madrid.

Dubois, P. du (2021) *Trojan horse. Saving the classics from conservatives*. New York: NYU Press.

Entrialgo, M. (2024) *Malismo: La ostentación del mal como propaganda*. Madrid: Capitán Swing.

Español, D. y Franco, J. (2021) *Recreación histórica y didáctica del patrimonio: nuevos horizontes para un cambio de modelo en la difusión del pasado*. Gijón: Ediciones Trea.

Fabien van Geert, F. (2020) *Du musée ethnographique au musée multiculturel. Chronique d'une transformation globale*. Paris: La Documentation Française.

Gifford-Gonzalez, D. (1993) 'You can hide, but you can't run: Representation of women's work in illustrations of palaeolithic life', *Visual Anthropology Review*, 9, pp. 3–21.

González Gutiérrez, P. (2018) 'Cuando solo nos queda enseñar las cartas: sesgos androcéntricos en el análisis de la vida de Trota de Ruggiero', *Nomadías*, 25(1), pp. 55–67.

González Gutiérrez, P. (2021) *Soror. Mujeres en Roma*. Madrid: Desperta Ferro.

González Gutiérrez, P. (2025) *¿Existieron las romanas?* Madrid: Akal.

Grever, M. y Van Nieuwenhuyse, K. (2020) 'Popular uses of violent pasts and historical thinking', *Journal for the Study of Education and Development*, 43(3), pp. 483–500.

Haraway, D. (1984) 'A cyborg manifesto', *Socialist Review*, 15, pp. 65–107.

Hernández Corchete, S. (2008) *La historia contada en televisión: el documental televisivo de documentación histórica en España*. Barcelona: Gedisa.

Hodkinson, S. (2022) 'Spartans on the Capitol. Recent far-right appropriations of Spartan militarism in the USA and their historical roots', in Beerden, K. and Epping, T. (eds.) *Classical controversies: reception of Graeco-Roman antiquity in the twenty-first century*, pp. 59–84. Leiden: Sidestone Press.

Hornos, F. y Rísquez, C. (2000) 'Paseando por un museo y buscando el lugar de la mujer', *Arqueología espacial*, 22, pp. 175–186.

Mantilla, K. (2015) *Gendertrolling: How misogyny went viral*. Santa Barbara: Praeger.

Naerebout, F. G. (2022) 'Introduction: stop the steal!', in Beerden, K. and Epping, T. (eds.) *Classical controversies: reception of Graeco-Roman antiquity in the twenty-first century*, pp. 59–84. Leiden: Sidestone Press.

O'Connor, J. E. (2001) 'History in images/images in history: Reflections on the importance of film and TV study for an understanding of the past', *The American Historical Review*, 93(5), pp. 1200–1209.

Paz, M. A. (2015) 'Usos públicos de la historia en la Transición española. Divulgación histórica y debate en Televisión Española (1978 a 1985)', *Historia y política: ideas, procesos y movimientos sociales*, 33, pp. 275–302.

Piñeiro Otero, T. et al. (2024) '¿Sueñan los troles con mujeres en el poder? Una aproximación al troleo de género como violencia política', *Anuario electrónico de estudios en comunicación social "Disertaciones"*, 17(2), a.13988.

Querol, M. A. (2013) 'Las mujeres en los discursos y representaciones de la Prehistoria: una visión crítica', in Domínguez, A. (ed.) *Política y género en la propaganda en la Antigüedad: antecedentes y legado*, pp. 63–80. Gijón: Trea.

Sánchez Romero, M. (2022) *Prehistoria de mujeres*. Madrid: Destino.

Segura Graíño, C. (2001) 'Actividades remuneradas y no remuneradas de las mujeres en la Edad Media hispana', in V.V.A.A. (eds.) *Aragón en la Edad Media: rentas, producción y consumo en España en la baja Edad Media. Sesiones de trabajo*, pp. 109–120. Zaragoza: Universidad de Zaragoza.

Skinner, M. (1987) 'Classical studies, patriarchy and feminism: The view from 1986', *Women's Studies International Forum*, 10, pp. 181–186.

Stone, L. (1985) 'Only women', *The New York Review*. Available at: https://www.nybooks.com/articles/1985/04/11/only-women/ (Accessed: 23 December 2025).

Venegas, A. (2023) *Pantallas de la memoria. Cómo y por qué las imágenes digitales transforman nuestra idea de la historia*. Madrid: Clave Intelectual.

Zuckerberg, D. (2019) *Not all dead white men. Classics and misogyny in the digital age*. Cambridge, MA: Harvard University Press.

RECURSOS WEB

EXIST Project (2021) 'Resultados del proyecto'.
Disponible en: https://nlp.uned.es/exist2021/ (Consultado: 23 de diciembre de 2025).

Independent (2019) 'Paul Gauguin, National Gallery, #MeToo…'.
Disponible en: https://www.independent.co.uk/arts-entertainment/paul-gauguin-national-gallery-me-too-art-harassment-assault-a9216801.html (Consultado: 23 de diciembre de 2025).

PSMag (2013) 'Women aren't welcome on the internet'.
Disponible en: https://psmag.com/social-justice/women-arent-welcome-internet-72170/ (Consultado: 23 de diciembre de 2025).

The Guardian (2014) 'Mary Beard: vocal women treated as freakish androgynes'.
Disponible en: https://www.theguardian.com/world/2014/feb/14/mary-beard-vocal-women-treated-freakish-androgynes (Consultado: 23 de diciembre de 2025).

Wondermark (s. f.) 'Comic #1062'.
Disponible en: https://wondermark.com/c/1062/ (Consultado: 23 de diciembre de 2025).

YouTube (s. f.) '[Vídeo]'.
Disponible en:
https://www.youtube.com/watch?v=yRNt7ZLY0Kc (Consultado: 23 de diciembre de 2025).

Hacer eterna a Roma: la divulgación como viaje compartido

Making Rome Eternal: Public Outreach as a Shared Journey

Iban Martín Prado
Creador de *Roma Aeterna*
iban.martin@gmail.com
ⓘ 0009-0008-4071-3116

1. INTRODUCCIÓN: EL RETORNO DE LA ORALIDAD EN LA ERA DIGITAL

La divulgación histórica atraviesa un momento de redefinición crítica en el ecosistema digital. Habitamos una "economía de la atención" tiranizada por la imagen, donde plataformas como *TikTok* o *Instagram* premian el impacto visual efímero y la fragmentación del mensaje. En este contexto de ruido visual, el formato audio emerge con una propuesta de valor radicalmente distinta: la recuperación de la intimidad y la pausa.

Este artículo disecciona, desde la praxis del proyecto *Roma Aeterna*, cómo la narrativa sonora permite alcanzar una profundidad analítica a menudo vetada en otros formatos. El podcast representa hoy una "hoguera digital", un espacio de resistencia que conecta directamente con la tradición oral primigenia. La ausencia de estímulo visual, lejos de ser una carencia, funciona como un catalizador cognitivo: obliga al oyente a construir activamente la imagen mental, favoreciendo una retención mnemotécnica superior a la del espectador pasivo.

El objetivo de este texto consiste en analizar los desafíos metodológicos de trasladar el rigor historiográfico a este lenguaje. Se explora cómo convertir la investigación académica en un relato de masas, demostrando que la complejidad, cuando se traduce con pasión y método, encuentra un lugar legítimo y sostenible en el mercado cultural contemporáneo.

2. EL ORIGEN DE *ROMA AETERNA*

La creación de *Roma Aeterna* es un ejercicio de resistencia cultural, una apuesta personal que, al analizarse bajo la lupa académica, revela las tensiones profundas entre la creación de contenido y las demandas del mercado.

Un proyecto cultural tan marcado y acotado implica una renuncia importante. La decisión estratégica de narrar cronológicamente la historia de Roma *Ab Urbe Condita* (desde su fundación) implicó renunciar conscientemente al *mainstream*, a la viralidad inmediata. El algoritmo de las plataformas de *streaming* impone lo que

podríamos llamar la "dictadura del clip": exige videos breves, descontextualizados y centrados en figuras potentes como Cayo Julio César o episodios que se viralicen rápidamente.

Sin embargo, traicionar la cronología hubiera sido traicionar la esencia del proyecto. Optar por la linealidad absoluta es un riesgo comercial, pero una necesidad intelectual: obliga a la audiencia a transitar por periodos oscuros y áridos, creando por el camino un vínculo con el oyente. Mi experiencia demuestra que, contra todo pronóstico, subestimamos al público: existe una audiencia hambrienta de contexto que agradece la densidad narrativa frente a la superficialidad imperante. El pasado de los personajes más importantes de Roma se convierte también en el pasado del oyente. Cuando César, siguiendo el ejemplo de una figura potente, se erija como heredero de Cayo Mario entenderá el contexto porque habrá vivido el pasado.

La elección del formato audio responde a una búsqueda de intimidad que la imagen a menudo destruye. El vídeo exige una atención tiránica; el audio, en cambio, acompaña. Desde su explosión en España en el año 2020, el podcast se ha convertido en un medio en auge. Un medio que concibo como una "hoguera digital": un retorno a la tradición oral primigenia, donde la voz del narrador es solo la mitad de la ecuación.

Esta es la paradoja del medio: al eliminar el estímulo visual, obligamos al oyente a una participación. Yo narro la historia, mientras el oyente construye la cabaña del Palatino en su mente. Esta co-creación imaginativa genera un vínculo emocional mucho más sólido que el de la pantalla.

La realidad del divulgador histórico dista mucho del romanticismo que se le atribuye; es un oficio solitario, de encierro y biblioteca. Mi metodología de trabajo se basa en el diálogo directo con las fuentes primarias y la investigación actual, un proceso que a menudo revela contradicciones insalvables entre autores como Tito Livio y Polibio y entre los autores y la realidad.

A pesar de que la tentación narrativa siempre es ofrecer un relato de película, coherente y cerrado, la ética de la divulgación reside en hacer justo lo contrario: exponer la contradicción. En lugar de ocultar la duda, la convierto en protagonista de la trama. Explicar al oyente que no sabemos algo o el por qué una fuente miente o exagera es, quizás, el acto pedagógico más potente del proyecto. Busco que el público aprenda a navegar la incertidumbre de la historia y poner en duda relatos que se han dado por sentados durante mucho tiempo.

Finalmente, es necesario abordar la dimensión afectiva de este trabajo. Transformar el dato académico frío en una narración emocionante es un proceso de traducción constante. ¿Cómo contar la crueldad antigua sin juzgarla, pero sin blanquearla?

El divulgador se enfrenta al micrófono en soledad, sin el *feedback* inmediato de un aula presencial, lanzando mensajes al vacío con la esperanza de que resuenen. Esta responsabilidad pesa: saber que, al otro lado, miles de personas están integrando esa narrativa en su visión del mundo obliga a un rigor autoimpuesto, a un compromiso de honestidad con la comunidad que escucha.

3. DESAFÍOS DE LA DIVULGACIÓN

Una vez establecida la conexión con el oyente, el divulgador se topa con un muro invisible pero sólido: el imaginario popular preexistente. Mi experiencia con *Roma Aeterna* revela tres frentes de conflicto principales donde el rigor académico choca con la cultura de masas.

El primer choque es sensorial. La audiencia llega al podcast con una Roma cinematográfica instalada en el córtex: escenarios de mármol blanco impoluto, saludos con el brazo derecho en alto y gladiadores peleando contra rinocerontes. Como divulgador, mi trabajo sucio consiste en dinamitar esa imagen estética sin que el oyente pierda la fascinación por el periodo.

Es un proceso que genera lo que en psicología se llama disonancia cognitiva. Cuando describo una Roma policromada con colores chillones, que huele a *garum* y letrina, y que es ruidosa y caótica, noto una resistencia casi física en el público. La realidad arqueológica es a menudo decepcionante frente a la fantasía estética.

He aprendido que la verdad histórica es más difícil de "vender" porque es, esencialmente, más fea y desordenada que la ficción. El reto no es solo corregir un dato, sino sustituir todo un escenario mental confortable por una realidad incómoda que a veces genera rechazo.

Más profunda que la resistencia estética es la resistencia ética. El público general busca instintivamente narrativas maniqueas de "buenos y malos". Sin embargo, al descender a las fuentes primarias, la escala de grises es la única constante. Aquí es donde la divulgación histórica puede activar el rechazo que ya se había generado al presentar la realidad incómoda.

Humanizar a las figuras históricas y aportar realidad a relatos fijados implica, inevitablemente, exponer sus miserias y contradicciones. He recibido correos de oyentes genuinamente molestos solo por hablar de mujeres y sexualidad sin los filtros caducos de siempre. Existe una tensión constante entre el deseo del público de idealizar el pasado y mi obligación de contextualizarlo. Mi postura es firme: dulcificar la historia para no molestar es convertirla en una caricatura. La función de la divulgación es combatir el presentismo, enseñando a la audiencia a no juzgar a los romanos con la moral del siglo XXI, sino a comprenderlos en su contexto.

Finalmente, el tercer gran reto es puramente lingüístico. ¿Cómo hablar de *potestas* o *auctoritas* a una persona que te escucha a las siete de la mañana en un

vagón de metro sin que se duerma? En *Roma Aeterna*, he desarrollado un estilo que denomino "coloquial culto". Integro analogías de la cultura pop moderna (comparar la política senatorial con *Juego de Tronos* o dinámicas parlamentarias actuales) estrictamente como puentes cognitivos, nunca como sustitutos del concepto. También el humor juega aquí un papel vital, aunque es peligroso. El chiste sirve para liberar tensión tras un bloque de datos denso, para dejar que el cerebro "respire", pero nunca puede ser el protagonista. Si el oyente recuerda la gracia, pero olvida el hecho histórico, considero que el episodio ha fallado.

En los últimos años, un nuevo desafío se ha sumado a la ecuación: la irrupción de la Inteligencia Artificial generativa. Hoy en día, cualquier creador puede generar en segundos una imagen hiperrealista de una legión romana con una estética impactante, perfecta para las redes sociales.

Sin embargo, para un proyecto de rigor histórico, estas herramientas representan una trampa. Las IAs inventan arquitecturas fantásticas y perpetúan estereotipos visuales erróneos. Utilizarlas sería traicionar la promesa básica del podcast: la veracidad. Si dedico 50 minutos a desmontar mitos con la voz, no puedo reforzar esos mismos mitos con la carátula del episodio o con un espectacularmente falso video promocional.

Esta postura ética conlleva una clara desventaja competitiva. Al renunciar a la espectacularidad de la imagen generada por IA, mis contenidos compiten visualmente en inferioridad de condiciones contra canales que sí la utilizan. Es un hándicap autoimpuesto: el algoritmo premia la imagen impactante (aunque sea falsa), mientras que el rigor exige a menudo imágenes de restos arqueológicos reales o monedas, que son objetivamente menos atractivos para el ojo no entrenado. Asumir esta pérdida de alcance es el precio a pagar por no contribuir a la polución visual del imaginario histórico.

4. SOSTENIBILIDAD Y COMUNIDAD

La validación final de cualquier proyecto de divulgación reside en su impacto real y su viabilidad. En el caso de *Roma Aeterna*, la transición de lo digital a lo físico y la estructura de financiación revelan claves importantes sobre el consumo cultural contemporáneo.

Durante la primera etapa del proyecto (2020-2023), mi relación con la audiencia era puramente digital: estadísticas de descargas, comentarios en los programas y redes sociales y correos electrónicos. Sin embargo, a partir de 2023 decidí dar un paso más: organizar viajes culturales a Roma, lo que transformó radicalmente mi comprensión del formato.

Es en el terreno, frente a la evidencia arqueológica, donde se manifiesta la potencia del audio. Ahí descubrí el poder del vínculo con el oyente. El turista convencional, al observar el Foro Romano sin contexto, ve las ruinas, quizá lee algo

al respecto y sigue adelante. El oyente del podcast, en cambio, proyecta sobre esas ruinas la narrativa construida durante horas de escucha. Como decía antes, se ha convertido en parte de su pasado.

No necesitan ver una reconstrucción 3D en una pantalla; la imagen ya está instalada en su memoria. Cuando señalo el Templo de Cástor y Pollux, señalo el emplazamiento del *Comitium* o el Templo del Divino Julio, se activa un conocimiento latente. El podcast dota a la visita de significado previo y potencia su disfrute.

Desde el principio soy consciente de que este proyecto desafía la lógica del mercado publicitario tradicional. Si *Roma Aeterna* dependiera exclusivamente de la publicidad programática o de las visitas masivas, no sería viable, ya que también apuesto por alejarme de publicidad dañina: casas de apuestas, bancos, inmobiliarias… que podrían financiar con publicidad el proyecto pero que lo harían perder su identidad.

Mi modelo es de puro mecenazgo. Al igual que Cayo Mecenas financiaba a Horacio o Virgilio para liberarles de la necesidad material, hoy son cientos de micro-mecenas quienes sostienen el proyecto. Esto tiene también una implicación editorial directa: mi único jefe es la comunidad.

Eso me ha llevado a descubrir una verdad que ha confirmado el objetivo inicial del proyecto: el mecenas paga exclusivamente por mantener ese rigor. La financiación directa compra mi tiempo para investigar, leer tres bibliografías contradictorias y ofrecer un producto cuidado. Se establece así un círculo virtuoso: la comunidad financia la independencia, y esa independencia garantiza la calidad del contenido.

Roma Aeterna demuestra que la divulgación histórica honesta y rigurosa es viable económicamente si se cuida el vínculo de confianza con la audiencia, rompiendo el mito de que en internet solo funciona lo superficial.

5. Conclusiones: el puente hacia la lectura y la eternidad

La trayectoria de *Roma Aeterna* permite extraer conclusiones que trascienden la anécdota personal y apuntan hacia nuevas vías para la divulgación histórica.

En primer lugar, se refuta la idea de que la divulgación digital compite con el saber académico. Al contrario, mi experiencia demuestra que el podcast actúa como una puerta de entrada. Lejos de saciar la curiosidad con un producto cerrado, el formato audio la inflama. Recibo constantemente testimonios de oyentes que, tras escuchar un programa, acuden a las librerías en busca de la *Farsalia* de Lucano o los *Comentarios* de César, o incluso los hay que se han matriculado a estudiar Historia. El podcast opera, por tanto, como un puente de alfabetización histórica: entrenamos al público para perder el miedo a la complejidad, creando una nueva base de lectores

preparada para ya sea consumir la producción académica universitaria o generarla tras estudiar.

Finalmente, este proyecto me ha llevado a redefinir el propio concepto que le da título. Si observamos la evidencia material, la ciudad no es eterna; las estructuras colapsan y el mármol se recicla, pasando de un edificio a otro. La eternidad, por tanto, reside fuera de la arquitectura.

La antigüedad perdura únicamente porque mantenemos activa la conversación sobre ella. Cada vez que divulgamos con rigor, cada vez que explicamos las contradicciones de Augusto o la vida cotidiana de las personas más desfavorecidas, estamos inyectando vida a esa memoria. Nosotros somos el soporte vital. La historia se mantiene viva, precisamente, al contarla con todos sus matices, sus grietas y su complejidad radical. Esa es la manera de hacer eterna a Roma

REFERENCIAS BIBLIOGRÁFICAS DE INTERÉS

Aguado, O. (2020) 'Roma victor vs. Roma victrix: Recepción cinematográfico-televisiva de la guerra y el ejército romano en el siglo XXI'. Disponible en: https://www.academia.edu/42385259 (Consultado: 25 de febrero de 2026).

Aparicio Resco, P. (2025) *Ventanas al pasado: La reconstrucción de la historia con imágenes*. Madrid: Akal.

Carrión, J. (2020) *Lo viral*. Barcelona: Galaxia Gutenberg.

Espinosa de los Monteros, M. J. (2020) 'El imparable auge del podcast', *El País Semanal*, 11 October. Disponible en: https://elpais.com/elpais/2020/10/09/eps/1602258181_939048.html (Consultado: 25 de febrero de 2026).

iVoox (2025) *Observatorio iVoox 2025: Estado del podcast en español*. Disponible en: https://landings.ivoox.com/observatorio2025 (Consultado: 25 de febrero de 2026).

Pons, A. (2013) *El desorden digital: Guía para historiadores y humanistas*. Madrid: Siglo XXI.

Sánchez Marcos, F. (2012) *Las huellas del futuro: Historiografía y cultura histórica en el siglo XX*. Barcelona: Universidad de Barcelona.