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,

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 Archivist*, 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).