# Memory Without Origin

*Why Research Libraries and Archives Need Governance Infrastructure for the AI Training Era*

Leo S. Lo

April 2026

*Current AI foundation model training permanently* ***eliminates an institution's ability to govern how its collections are used unless contractual protections are established in advance. It is a stewardship, legal, and ethical decision, not an access decision, and it cannot be undone.*** *This paper explains why that distinction matters, what it costs the profession to ignore it, what it means for the communities and donors whose consent underpins archival trust, and why the UVA Archival AI Protocol was developed as a governance response. It does not argue for restricting who can use archival collections. It argues for governing what institutions do with them.*

## Memory Without Origin

*Imagine a world where you know things but cannot say where you learned them.*

You know that a particular community's land was taken in 1887. You know the details of a negotiation, the names of the parties, the terms that were agreed and the terms that were broken. You know this with confidence. But you cannot point to the record that established it. You cannot say which archive held the documents, which collection they came from, which community entrusted them to which institution under what conditions. The knowledge circulates. The origin does not.

Imagine a researcher who receives an answer from an AI system about the contents of a historical collection. The answer is specific. It draws on real materials. But the researcher cannot verify which items were consulted, cannot check whether the context was preserved, cannot determine whether the community whose records were used ever consented to this use. The answer exists. Its provenance does not.

This is not a hypothetical future. It is the predictable consequence of a specific technical act conducted at scale: absorbing archival materials into foundation model weights without provenance conditions. The knowledge enters circulation permanently. The chain of custody that archivists have maintained for generations breaks at the moment of ingestion. What remains is memory without origin: authoritative-seeming outputs whose archival sources cannot be identified, verified, corrected, or contested.

Archives exist because society decided that knowledge with provenance is more trustworthy than knowledge without it. The archivist's core professional function is maintaining the connection between records and their origins, so that future users can evaluate, contest, and build on the past with full awareness of where the evidence came from and under what conditions it was created.

AI training requests are arriving now, the terms are being set now, and the consequences of getting them wrong are more permanent than anything the profession has previously encountered. Once archival materials are absorbed into foundation model weights, no subsequent

institutional action can remove them from the model. The UVA Archival AI Protocol was developed to address that problem before the window closes.

This paper addresses a specific context: archival and special collections where materials are sensitive, donor-bound, or community-linked - collections entrusted under conditions that predate and cannot anticipate commercial AI training. General circulating collections and widely held published works raise different questions under different frameworks. The governance argument here is not a claim about library access norms broadly. It is a claim about what custodial ethics require when irreversible AI training would breach existing obligations or permanently foreclose future ones.

## A Stewardship Decision, Not an Access Decision

The commitment to broad, nondiscriminatory access to knowledge is a genuine and foundational library and archival value. The debate about AI training governance in libraries and archives is often framed as a test of that commitment: should libraries restrict who can use their collections and for what purposes? That framing is understandable. It is also importing the wrong professional doctrine for the situation at hand.

Access doctrine asks: may this user consult these materials? The answer defaults to yes absent specific identifiable harm. Stewardship doctrine asks: may this institution take a permanent, consequential action with these collections? The answer requires affirmative professional justification. These doctrines exist for principled reasons and have always been treated differently in professional practice.[1]

Archivists make stewardship decisions every day. Honoring a donor restriction is a stewardship decision. Declining to deaccession materials under institutional pressure is a stewardship decision. Complying with NAGPRA repatriation requirements is a stewardship decision. None of these diminish the commitment to access. They reflect the profession's recognition that some institutional acts have consequences that outlast any single use, and require a different standard of justification.

AI foundation model training is a stewardship, legal, and ethical decision, not an access decision. A researcher reading an oral history transcript accesses the collection. A commercial AI pipeline ingesting that transcript into model weights permanently changes the institution's relationship with the collection and with the communities whose records it holds. Those weights will

[1]SAA Core Values state that archivists "should promote and provide the widest possible access to materials, while respecting legal and ethical access restrictions including public statutes, cultural protections, donor contracts, and privacy considerations," and affirm that access may be justifiably limited in some instances while archivists still seek to foster open access and unrestricted use as broadly as possible when appropriate. Society of American Archivists, Core Values of Archivists (2020). The SAA Code of Ethics reinforces this balance, stating that archivists "actively promote open and equitable access to records in their care" while working with "creators, donors, organizations, and communities to ensure that any restrictions applied are appropriate, well-documented, and equitably enforced." Society of American Archivists, Code of Ethics for Archivists (2012, reaffirmed 2020).

influence millions of future interactions in products the institution cannot inspect. The decision cannot be undone regardless of what the institution learns afterward.

After training completes, no future archivist can make a different decision about those collections' influence on that system. The reversibility principle is foundational to archival professional standards: actions taken on materials should be reversible. That principle draws a line not at physical permanence but at the foreclosure of future professional choice. Deaccessioning removes materials from circulation. Repatriation returns them to their communities of origin. Both acts resolve the institutional relationship with the materials. AI training does the opposite: it multiplies the materials invisibly across model weights, generating ongoing influence over future knowledge production that no subsequent institutional decision can audit, correct, or recall.

The Protocol governs one specific act: the permanent absorption of collection materials into foundation model weights where provenance disappears and institutional governance cannot follow. It explicitly permits retrieval-based systems, which index materials so they can be retrieved with citations at query time and removed if circumstances change. It permits computational research under institutional oversight and narrow internal organizational uses. Researchers, students, and the public can read, download, cite, teach with, and build on these materials without restriction.

Retrieval-based systems and foundation model training are not alternative architectures competing to solve the same problem. They are categorically different acts. With current retrieval-based architectures, the institution knows what was retrieved, can cite it at item level, and can remove it if circumstances change - the provenance chain holds. Foundation model training breaks that chain. Knowledge enters the model. Its origins do not follow it in.

Provenance requirements operate at two distinct stages. Input logging creates a dataset-level record of which materials entered training - achievable with current technology, and essential for documenting institutional participation, invoking contractual rights, and establishing accountability for benefit-sharing. Output citation, tracing a generated response back to a specific source document, is achievable with current retrieval-based architectures and is fully specified in the Protocol's Appendix B. For foundation model training, tracing outputs back through model parameters to specific training documents produces, with current training methods, probabilistic approximations rather than archival provenance.

*That is a genuine technical limitation.*

It is also the reason the Protocol sets the provenance standard where it does. When an act is irreversible, the governance standard must be higher than what current capability can satisfy, not lower. A standard calibrated to what today's technology can deliver would permit decisions that cannot be undone based on assurances that cannot be verified. The Protocol declines to do that. Institutions that cannot meet the provenance standard should decline foundation model training requests under the Protocol's presumption against approval. That is not the standard failing. That is the standard working - and meeting the ethical codes and living professional values the archival field has established.

As interpretability research advances, the technical capacity for output attribution may improve. If it does, the Protocol should evolve to reflect what can then be verified. A standard set against

current limitations is not a permanent ceiling - it is the minimum the profession can responsibly accept until the technology catches up to the obligation.

## Community and Donor Obligations

The SAA Core Values establish that archivists bear affirmative obligations to donors, records creators, and communities whose materials were entrusted to institutions under specific conditions. The document explicitly names donor contracts and cultural protections as legitimate bases for access restrictions, and describes materials throughout as entrusted to archivists, not owned by them. The SAA Code of Ethics states directly that archivists actively promote open and equitable access to records in their care, strive to minimize restrictions, and work with creators, donors, organizations, and communities to ensure that any restrictions applied are appropriate, well-documented, and equitably enforced - including that material designated by communities as culturally sensitive should not be disseminated or reproduced without express permission from the communities of origin. That is a use-based access condition stated without qualification in the profession's own ethics document. A governance framework for AI training uses of archival collections is not a departure from archival professional values. It is an application of those values to a situation the profession has not previously faced in this specific form.

The CARE Principles for Indigenous Data Governance establish that communities have governance rights over their own records that exist independently of institutional ownership and independently of copyright. For collections with active community relationships, bulk AI training ingestion without community consultation is not an access question. It is a question about the custodial relationship under which materials were entrusted to the institution. The memory without origin problem is most acute for exactly these collections. Oral histories ingested into a commercial AI model without community consent do not simply lose attribution. They enter a system that will generate outputs drawing on those voices, that those communities cannot audit, correct, or remove.

*That breach of the custodial relationship is permanent.*

## What the Protocol Requires and Why

The communities, donors, and researchers described in this paper are why the UVA Archival AI Protocol exists. Its core rule is straightforward: no access without accountability. It is organized around three pillars: provenance and attribution, donor and community responsibilities, and institutional stewardship. These are the minimum conditions for preventing memory without origin, honoring the custodial obligations under which collections were acquired, and preserving the institution's ability to govern its collections after an agreement is signed. The Protocol was developed at UVA and is offered as a contribution to the profession's collective effort to establish shared standards - not as a UVA position, but as a starting point the field can adopt, adapt, and improve.

An institution that removes materials because a donor restriction surfaces, corrects errors in how its collections are represented, or stops further use because a community has withdrawn consent is doing custodial work. That is what "accountability" means in the Protocol: not ownership or

market leverage, but the ongoing capacity to act on behalf of the communities and donors whose trust the institution holds.

The most common scenario where this accountability is at risk is the digitization-for-training exchange. Digitization and AI training are not the same act and should not be treated as one transaction. AI training absorbs content permanently into model weights. When a vendor proposes to fund digitization in exchange for AI training rights, the two must be evaluated separately. Otherwise short-term backlog relief quietly purchases long-term training rights whose value compounds in ways a one-time payment cannot track.

That scenario requires a vendor to surface the question. A second, less visible scenario requires nothing except a standard agreement signed without asking what the retained copy can become. A digitization agreement that grants the digitizing party retained copies with unconstrained downstream use rights has already resolved the AI training question - by default, in the vendor's favor, before the institution recognized the question was being asked. The provenance chain breaks not when the model trains, but when the institution ships the boxes. Digitized materials entering shared institutional repositories carry the same risk: the governance terms of the repository, not the original digitization agreement, determine what computational uses those materials ultimately permit.

Shared professional standards do practical work at the negotiating table - including negotiations that institutions do not recognize as AI training negotiations until it is too late. Vendors currently know what terms every institution has accepted. Institutions do not. A shared framework reduces that asymmetry, aligns counsel, procurement, archives, and library leadership around a common baseline, and establishes the floor that makes responsible partnerships possible rather than preventing them. For library and archive leaders who do not hold final authority over these decisions, a shared professional standard is also the most durable instrument available: it names the obligation, documents the consensus, and gives institutional leaders something to bring to the table with general counsel, provosts, and technology transfer offices that individual judgment alone cannot provide.

The tools exist. The window is closing. The decision is now.

*Leo S. Lo is University Librarian and Dean of Libraries at the University of Virginia and Special Advisor to the Provost on AI Literacy. He is past president of the Association of College and Research Libraries (2024–2025). The UVA Archival AI Protocol is available at https://doi.org/10.18130/5dqf-9w86 ; and the Adoption Kit (https://doi.org/10.18130/jbeg-a995) is available at the same repository.*