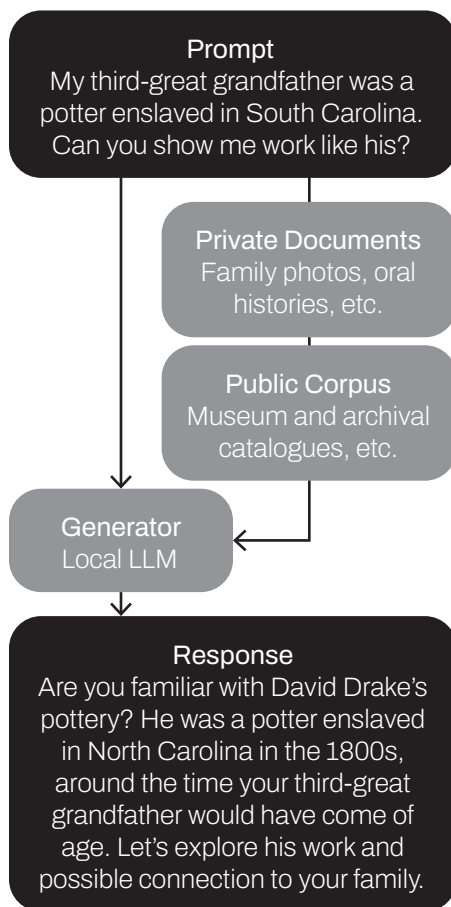


AI FOR REPAIR: EPISTEMIC JUSTICE, REPARATIVE TECH, AND THE LEGACY RECLAMATION AGENT



EXECUTIVE SUMMARY

This essay advocates for a Legacy Reclamation Agent (LRA). An LRA is an Artificial Intelligence (AI) agent that combines a local Large Language Model (LLM) and two-pronged Retrieval Augmented Generation (RAG) process to find correlations between privately compiled archives and a public corpus of online museum, archive, and auction records, then guides users through the process of seeking financial, cultural, and emotional restitution for their reclaimed legacies.

The issue is bigger than one family. Many people are disconnected from their ancestors because of generationally multiplied dispossession—an enduring economic, social, and psychological impact, in the U.S. context, of enslavement and colonization. Redressing this dispossession at scale is a pressing matter of social justice. Further, many people are not able to reconnect with their ancestors because navigating the systems that have captured knowledge about their lives and

work requires significant disciplinary expertise. Addressing this systemic deprivation is critical to cultivating a cultural sector that values epistemic justice.

The LRA puts the power of disciplinary experts into the hands of any descendant, rewriting a needle-in-a-billion-haystacks problem that relies on improbable human effort and luck into a navigable map to repair. It creates opportunities for rightful beneficiaries of lost legacies to find financial, cultural, and emotional restitution, and to claim narrative agency.

Perhaps most importantly, the LRA serves as a concrete model for reparative technology, setting a precedent for a new kind of socially responsible computing. AI is often lauded for making our lives easier by automating tasks off our plates and reducing friction. Underlying this essay is a conviction that instead of simply making our lives easier, AI could serve as an enabling technology that pulls together disparate information and creates the friction that is critical to cultivating a more just and equitable society.



Figure 01 Descendants of David Drake at the historic restitution ceremony. (Source: Museum of Fine Arts, Boston)



Figure 02 Screenshot of a news article titled, "Their enslaved ancestor's pottery sells for over \$1 million. They get nothing." (Source: The Washington Post)

DEFINING THE PROBLEM: LOST LEGACIES

I toggle back and forth between the Indian Land Grant PDF that I found with no context in a federal agency's online archive—I can't remember which one, I've queried so many—and the genealogical research platform. I hit ctrl+F, but I'm not even sure what to search the PDF for. I know that my great-grandmother was born somewhere near Albuquerque sometime around 1905 and that she was given some land by a public entity that led to familial unrest after her recent passing, but that's about all I know. Each of her granddaughters has a different story about how Memaw found her way to the Permian Basin, and her daughter has more contradicting stories than all of them combined.

In November of 2025, the Museum of Fine Arts, Boston (MFA), returned two monumental stoneware jars by enslaved artist David Drake to his known descendants. Mr. Drake made tens of thousands of ceramic vessels over the course of his life—many of which he risked severe punishment to sign, date, and inscribe couplets onto and all of which were presumably sold to benefit his enslaver, Lewis J. Miles. (MFA, 2025) About fifteen of Mr. Drake's descendants travelled to Boston for the restitution ceremony, which marked the first time that the MFA has resolved an ownership claim for works of art that were wrongfully taken under the conditions of slavery in the 19th-century U.S. (MFA, 2025)

Restitution, or "the return of cultural material to an individual, group or nation with the overall aim of doing justice for a past or ongoing wrong," is an increasingly prominent issue in archival and curatorial studies. (Laurin, 2022)

As historian Ciraj Rassool and anthropologist Victoria E. Gibbon write, "through curatorial practice, the deceased, possessions, plants, and animals were [...] removed from their communities [and] turned into museum objects." As a renewed politics of decolonization rises in heritage studies, "a major theme of decolonial enquiry into these unethical collecting legacies is the need for purposeful and proactive acts of return." (Rassool and Gibbon 2024)

Pauline Baker, Mr. Drake's great-great-great-granddaughter, described the MFA restitution ceremony as "a hallelujah moment." (NPR 2025) Her son, Yaba Baker, similarly called the return "a spiritual restoration" that gave the family back a piece of themselves. (Kindy 2025)

As emotionally resonant and spiritually fulfilling as restitution can be, though, reclaiming lost legacies can also have financial implications. Several years before the MFA restitution, an artwork by Mr. Drake sold for a reported \$1.56 million. Mr. Drake's descendants clearly ought to have benefited from this sale, but because no restitution had taken place in the case of that artwork, they did not—an injustice that highlights how dispossession can only multiply across generations absent repair. (Kindy 2025) The MFA restitution interrupts this dispossession and calls long-overdue public attention to the importance of ensuring that enslaved artists' descendants are the beneficiaries of those artists' legacy.

How might socially responsible computing aid in ensuring that rightful beneficiaries are reunited with lost legacies? To answer this question, we must rewind the clock to a 2016 call between Daisy Whitner, another of Mr. Drake's descendants, and genealogical

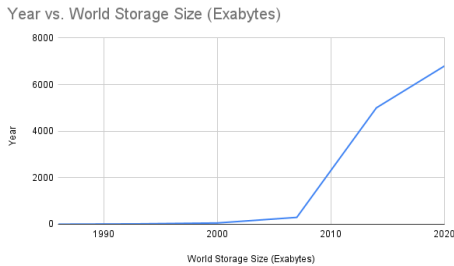


Figure 03 A chart showing the rapidly-increasing world storage size in exabytes. information management scholar Debin Xu writes, “[digitized] archives can use big data and cloud computing technology to [...] effectively solve the problem of a “bloated library” in traditional archives management.” (Xu 2022) Unfortunately, digital “bloated libraries” are still a major hurdle for the average person seeking to answer questions about their ancestry.

researcher April Hynes. Hynes came across Whitner’s name while conducting research on Mr. Drake’s life and decided to contact her with news of her “famous ancestor.” Until that phone call, the family had never heard of David Drake or his art. (Kindy 2023)

It isn’t that information on Mr. Drake’s life or work was not publicly available—on the contrary, an entire book on these topics was published in 2008 and 270 of Mr. Drake’s vessels are accounted for by various cultural institutions, many of which maintain ever-growing, publicly available electronic records. (Todd 2008; Burtseva 2023)

The problem is one of scale. There are hundreds of thousands of museums and archives worldwide. The U.S. National Archives alone boasts a permanent collection of more than 33 billion electronic records—if you spent just one second reviewing each one, it would take you more than one thousand years to get through them all. (National Archives, 2026) It’s easy to see how Mr. Drake’s 270 vessels, along with the legacy they comprise, could become flotsam in a sea of data—a sea of data only navigable with the guidance of curators and archivists, and with the genealogical path charted by Whitner and Hynes on their 2016 call.

The issue is bigger than one family. Many people are disconnected from their ancestors because of generationally multiplied dispossession—an enduring economic, social, and psychological impact, in the U.S. context, of enslavement and colonization. Redressing this dispossession at scale is a pressing matter of social justice. Further, many people are not able to reconnect with their ancestors because navigating the systems that have captured

knowledge about their lives and work requires significant disciplinary expertise. Addressing this systemic deprivation is critical to cultivating a cultural sector that values epistemic justice, or “the right of every people to their own knowledge and ways of generating, legitimizing and valuing it.” (Schmelkes 2023)

This calls for a new kind of assistive technology—a sector we might call “reparative technology.” Ancestry.com exists. Chatbots exist. Digital collections for archives, museums, and auction houses grow by the day. What we need now is a future-oriented tool that can synthesize existing information with a bias towards action and commitment to epistemic justice.

OUTLINING THE TOOL: THE LEGACY RECLAMATION AGENT

An activist friend points me towards the Legacy Reclamation Agent (LRA). The co-development guide, which has been crowd sourced and revised by the many people, families, and communities already using the local large language model-based tool, is easy enough to follow. I draft a list of guidelines loosely based on the publicly available Principles of Māori Data Sovereignty, both to hold myself accountable to the family members whose stories I’ll be stewarding and as a framework for any future collaborations that come out of the research. Then, I open the LRA and begin sharing what I know so far. I type the story of my great grandmother’s upbringing in New Mexico, adoption, and years in the Permian Basin. The LRA prompts me to expand, clarify, or upload documents based on what I’m sharing.

I begin recording my phone calls with my mom and aunts and uploading these as well. My grandmother’s sister sends me a big binder of

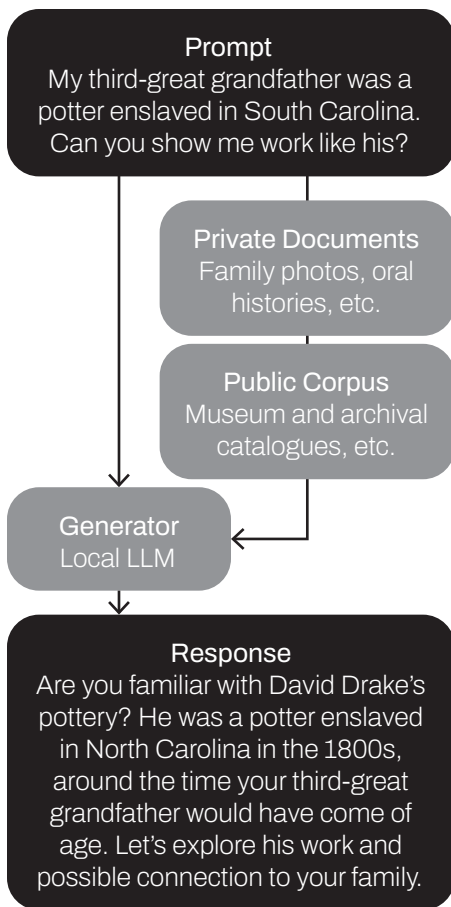


Figure 04 A basic diagram of the LRA's structure, combining a local LLM and two-pronged RAG process. (Source: author)

family photos to scan. Slowly, the LRA's database begins to populate with a virtual museum of my own family's history. Family portraits, oral histories, and scattered attempts at building out family trees in various online databases are supplemented by publicly available documents from federal agencies—this time, contextualized in layperson's terms so that we know what we're looking at. We finally have a coherent and communicable history.

The MFA restitution relied on disciplinary expertise and luck. The first (and so far, only) time that the MFA has resolved an ownership claim for works of art that were wrongfully taken under the conditions of slavery, that claim could not have been made if not for the long-shot connection between a passionate genealogical researcher and a receptive descendant of the artist.

What if technology could assist in writing a surefire script for restitution that didn't put rightful beneficiaries at the mercy of experts? Can AI support rightful beneficiaries as capable protagonists in demanding ownership of their legacies? Could legacy reclamation and restitution be a use case that bends the power of AI towards computing's social responsibility to repair?

This essay advocates for a Legacy Reclamation Agent (LRA). An LRA is an Artificial Intelligence (AI) agent that combines a local Large Language Model (LLM) and two-pronged Retrieval Augmented Generation (RAG) process to find correlations between privately compiled archives and a public corpus of online museum, archive, and auction records, then guides users through the process of seeking financial, cultural, and emotional restitution for their reclaimed legacies.

An LRA has two key technological components. The first is a two-pronged RAG process. RAG is the process that optimizes the output of an LLM by referencing a knowledge base outside of its training data before generating a response. In the LRA's case, that knowledge base holds two types of information: privately generated archives (oral histories, family photos, genealogy charts) and a continually indexed public corpus that grows as material is digitized by museums, archives, and auction houses.

RAG processes are proven to generate more targeted, reliable responses in LLMs. Commercial chatbots are prone to providing incomplete and occasionally hallucinatory answers. For example, when I ask ChatGPT to provide me with a comprehensive table of where David Drake's vessels are located today, it offers a list of 11 museums that account for a couple dozen jars, jugs, and pitchers—much fewer than 270—and insists that its list is comprehensive. Further, the tool sources first from news articles and Wikipedia pages as opposed to digital archives themselves, which is not ideal in formulating legally cognizable arguments for restitution. An RAG process addresses this issue.

The data archive produced as part of the LRA's RAG process is the property of its user. However, some stories are meant to be shared. Having a space to securely store, collate, and reflect on ancestral stories is an inherent benefit of the LRA. As appropriate, the digital data that is produced through the LRA's RAG process could also be instrumentalized to address algorithmic bias more generally, training more culturally informed and responsive models for the industry at large.



Figure 05 An illustration of Brooklyn Historical Society staff conducting an oral history interview. Oral histories are the kind of rich, intimate data that an RAG process could incorporate securely. (Source: Brooklyn Public Library)

The LRA's RAG process builds on a secure local LLM that users can engage with intuitively without compromising the privacy of intimate data. A secure local LLM can be just as intuitive as commercial chatbots, which hundreds of millions of people already interact with on a regular basis, while preserving data security. (Goth Itoi 2025) The model runs on hardware controlled by the user or a trusted partner, like a nonprofit law clinic or a tribal nation's IT department. Data—in this case, sensitive family histories, legal strategies, or community dialogues—never leaves that environment. Users can stop or modify the model's process without asking permission from a proprietary tech company, retaining sovereignty of the LRA.

A secure local LLM and thoughtful RAG process are crucial to a successful LRA, particularly as we consider the fraught relationship LLM-based tools have developed to profit, privacy, and transparency.

ADDRESSING THE RISKS: PROFIT, PRIVACY, AND TRANSPARENCY

As our family becomes more engaged in the project, my aunts start telling me stories “off the record.” Some of my younger cousins ask if using AI to work through these questions is a good idea with all the controversy that has come up around LLMs. Relatives who still live in New Mexico and maintain a deeper connection to the land are particularly incredulous, questioning who developed this technology and how they are turning a profit. I reassure them that the LRA is a sovereign tool just for us. We talk about the guidelines that I drafted at the beginning of the process, occasionally revising them to ensure that everyone is comfortable with how their data is used. I work with the

LRA to tag “off the record” stories in the RAG database so that they are only accessible to their author while still informing the arc of our family's history implicitly.

As AI chatbots permeate education, the workplace, and daily life, one does not have to look far to grasp the risks of LLM-based tools. For one, most tools are developed and operated by private companies ultimately seeking to turn a profit. These companies' commitment to the public interest is questionable, and they have often failed to prioritize users' data security—much less consider data sovereignty, or users' right to own and control their data. The commodification of user data is particularly insidious when those users are engaging the tool to redress dispossession. Further, black box models render these tools unintelligible and unchangeable—even if they garner buy in to begin with, which is unlikely.

The fact is that commercial LLM-based tools are not conducive to data sovereignty. These tools routinely ingest and centralize user inputs into U.S. or foreign servers, often implicitly claiming rights to use that data for training—a pragmatic nightmare for the retention of data sovereignty. A 2025 study at Stanford found that six U.S. frontier AI developers all employ their users' chat data to train and improve their models by default, and that some retain this data indefinitely. (King et al. 2025) This practice clearly compromises privacy and makes data readily commodifiable—particularly by multiproduct companies like Microsoft, for whom chatbot data can seep into how a user interacts with other products. (Goth Itoi 2025)

More fundamentally, while commitments to benefit “all of humanity” (taking OpenAI's charter

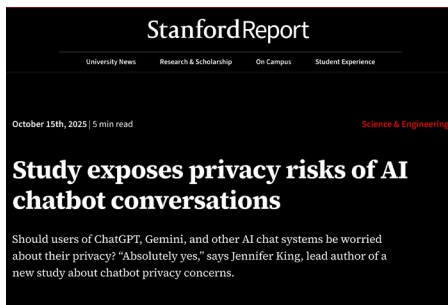


Figure 06 Screenshot of a news article titled, “Study exposes privacy risks of AI chatbot conversations,” and subtitled, “Should users of ChatGPT, Gemini, and other AI chat systems be worried about their privacy? “Absolutely yes,” says Jennifer King, lead author of a new study about chatbot privacy concerns.” (Source: StanfordReport)

as an example) may sound positive on-face, the principles of borderless, centralized data fusion that flatten epistemic difference and drive the development of commercial LLM-based tools produce an irreconcilable tension with users’ right to own and control their data. Ensuring data sovereignty is a part of cultivating an epistemically just reparative technology sector—which, as we have established, requires honoring the right of people to their own knowledge and ways of generating, legitimizing, and valuing it. As such, reparative technology will not thrive in a private sector that is largely unaccountable to the public interest its companies claim to serve.

This doesn’t mean that the fledgling sector needs to abandon LLM-based tools, though. Te Hiku Media, a small nonprofit Māori radio station in Aotearoa New Zealand, have made headlines for instrumentalizing AI to further their language revitalization work while centering data sovereignty. The group has achieved this largely through the development of a the Kaitiakitanga License, which details ground rules for collaboration. Kaitiakitanga is a Māori principle loosely translating to guardianship or stewardship, and the license only grants data access to organizations that pass benefits derived from its use back to the Māori people. (Hao 2022)

We can learn from the Te Hiku Media case and Kaitiakitanga License. Firstly, it isn’t enough to explain the LRA once it is developed. Rather, it is key that the development process is co-led, and ideally initiated by a would-be user. Further, it’s crucial that the LRA is owned by an actor that is both accountable to rightful beneficiaries and capable of sustaining a reclamation initiative, and that accountability is formalized through a license, list of principles,

or other document appropriate to the parties involved. Finally, even as the LRA is co-developed and owned by a formally accountable party, the development process itself needs to be clearly documented and repeatable by individuals or communities who might be interested in initiating a similar project.

A secure local LLM and thoughtfully structured RAG process are crucial to a successful LRA. This base structure can be further strengthened by implementing a responsible co-development process and an ownership model that formalizes accountability to center rightful beneficiaries’ data sovereignty. While users must have ultimate control over the privacy of the data that they produce, radical transparency in the architecture of the tool and its development process will ensure that the LRA can be picked up by as many people as possible to transform relationships between the past, present, and future.

THE GOAL OF THE GOAL: A PRECEDENT FOR REPARATIVE TECHNOLOGY

There are countless immediate use cases for the LRA. As an example, let’s return to David Drake. Since the MFA restitution, Mr. Drake’s descendants have established the Dave the Potter Legacy Trust. The trust seeks to “ensure [David Drake’s] descendants are among the beneficiaries of his legacy.” (Descendants of Dave 2026) A form and email address guide potential descendants through the process of claiming restitution. If the trust were able to co-develop, own, and use the LRA, it would be possible to build out a fuller picture of Mr. Drake’s blood descendants, and account for the full scope of his recapturable work, faster and more efficiently.

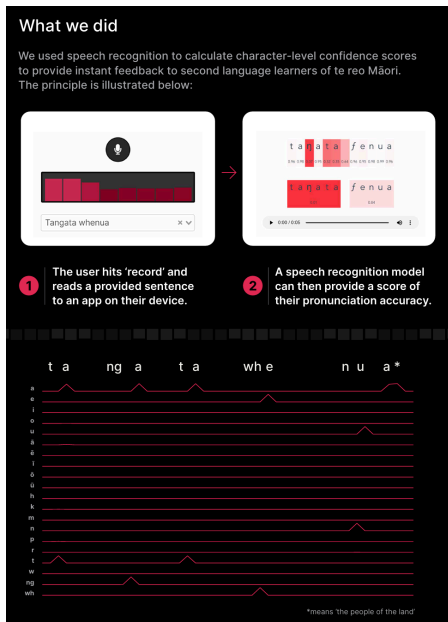


Figure 05 Clipping from a Te Hiku Media poster explaining how speech recognition was used to calculate character-level confidence scores to provide instant feedback to second language learners of te reo Māori. (Source: Te Hiku Media)

A secondary application for the LRA is as a proactive tool for institutions seeking to fulfill ethical mandates and audit their own collections—an increasingly prescient imperative. While it's crucial that reparative technology centers rightful beneficiaries as the primary use case, it's undeniable that a researcher like April Hynes could have located Mr. Drake's descendants much more quickly with assistance from the LRA, potentially allowing them to secure restitution earlier—an inarguable net benefit.

The LRA puts the power of disciplinary experts into the hands of any descendant, rewriting a needle-in-a-billion-haystacks problem that relies on improbable human effort and luck into a navigable map to repair. It creates opportunities for rightful beneficiaries of lost legacies to find financial, cultural, and emotional restitution, and to claim narrative agency.

Perhaps most importantly, the LRA serves as a concrete model for reparative technology, setting a precedent for a new kind of socially responsible computing. AI is often lauded for making our lives easier by automating tasks off our plates and reducing friction. Yet, anthropologist Anna Tsing reminds us that friction—“the awkward, unequal, unstable, and creative qualities of interconnection across difference”—is often what leads to “new arrangements of culture and power,” something we are in desperate need of today (Tsing 2004) Underlying the essay is a conviction that instead of simply making our lives easier, AI could serve as an enabling technology that pulls together disparate information and creates the friction that is critical to cultivating a more just and equitable society.

SOURCES

Bashar, M. A., Nayak, R., Knapman, G., Turnbull, P., & Fforde, C. (2023). An Informed Neural Network for Discovering Historical Documentation Assisting the Repatriation of Indigenous Ancestral Human Remains. *Social Science Computer Review*, 41(6), 2293–2317. <https://doi.org/10.1177/08944393231158788>

Burtseva, S. (2023, June 16). David Drake's Poetic Pottery Was Resistance. National Gallery of Art. <https://www.nga.gov/stories/articles/david-drakes-poetic-pottery-was-resistance>

Descendants of Dave. (n.d.). Dave the Potter Legacy Trust. Retrieved February 8, 2026, from <https://descendantsofdave.org/>

Goth Itoi, N. (2025, October 15). Study exposes privacy risks of AI chatbot conversations. Stanford Report. <https://news.stanford.edu/stories/2025/10/ai-chatbot-privacy-concerns-risks-research>

Hao, K. (n.d.). A new vision of artificial intelligence for the people. MIT Technology Review. Retrieved February 8, 2026, from <https://www.technologyreview.com/2022/04/22/1050394/artificial-intelligence-for-the-people/>

Kindy, D. (2023, April 2). Their enslaved ancestor's pottery sells for over \$1 million. They get nothing. The Washington Post. <https://www.washingtonpost.com/history/2023/04/02/dave-potter-enslaved-descendants/>

King, J., Klyman, K., Capstick, E., Saade, T., & Hsieh, V. (2025). User Privacy and Large Language Models: An Analysis of Frontier Developers' Privacy Policies (arXiv:2509.05382). arXiv. <https://>

doi.org/10.48550/arXiv.2509.05382

Krupa, K. L., & Grimm, K. T. (2021). Digital Repatriation as a Decolonizing Practice in the Archaeological Archive. *Across the Disciplines*, 18(1–2), 47–58. <https://doi.org/10.37514/ATD-J.2021.18.1-2.05>

Laurin, evlyne. (2022, February 2). Contemporary Art History 101–What is the Deal with Restitution? ..ART. <https://art.art/blog/contemporary-art-history-101-what-is-the-deal-with-restitution/>

Mahelona, K., Leoni, G., Duncan, S., & Thompson, M. (2023, January 24). OpenAI's Whisper is another case study in Colonisation. *Papa Reo*. <https://blog.papareo.nz/whisper-is-another-case-study-in-colonisation/>

Museum of Fine Arts, Boston. (2025, October 29). Museum of Fine Arts, Boston, Resolves Ownership of Works by Enslaved Artist David Drake. <https://www.mfa.org/press-release/david-drake-ownership-resolution>

National Archives. (2022, September 26). National Archives by the Numbers. National Archives. <https://www.archives.gov/about/info/national-archives-by-the-numbers>

OpenAI. (n.d.). OpenAI Charter. Retrieved February 8, 2026, from <https://openai.com/charter/>

Perel, M., & Elkin-Koren, N. (2022). Democratic Friction in Speech Governance by AI (SSRN Scholarly Paper No. 4293274). Social Science Research Network. <https://papers.ssrn.com/abstract=4293274>

Rassool, C., & Gibbon, V. E. (2024). Restitution versus repatriation: Terminology and concepts matter.

American Journal of Biological Anthropology, 184(1), e24889. <https://doi.org/10.1002/ajpa.24889>

Schmelkes, S. (2023, January 9). Epistemic justice and the knowledge commons for lifelong and lifewide. UNESCO. <https://www.unesco.org/en/articles/epistemic-justice-and-knowledge-commons-lifelong-and-lifewide-learning>

Shea, A. (2025, November 14). Boston museum returns stoneware jars made by enslaved man to his descendants. NPR. <https://www.npr.org/2025/11/14/nx-s1-5607225/boston-museum-returns-stoneware-jars-made-by-enslaved-man-to-his-descendants>

Te Hiku Media. (2023, October 11). Data Sovereignty and the Kaitiakitanga License. Te Hiku Media. <https://tehiku.nz/p.G75>

Todd, L. (2008). *Carolina Clay: The Life and Legend of the Slave Potter Dave*. W. W. Norton & Company.

Tsing, A. L. (2004). Introduction. In *Friction*.

Willingham, L. (2025, November 20). Descendants obtain works of enslaved potter in landmark restitution deal. AP News. <https://apnews.com/article/david-drake-pottery-slavery-museum-c6f3576b538ee7420766bfd38bde4bd8>

Xu, D. (2022). An Analysis of Archive Digitization in the Context of Big Data. *Mobile Information Systems*, 2022(1), 1517824. <https://doi.org/10.1155/2022/1517824>