

Tracing Borrowed Authority in Relational AI Systems

Hanjing Shi

Lehigh University
hasa23@lehigh.edu

Abstract

This abstract presents a focused branch of my broader research as a doctoral student: what must be visible when AI systems borrow authority from relational sources. I examine this question through two linked cases. The first is relational persistence: a public audit of AI afterlife and adjacent representation services, asking how products make grief, continuing bonds, social recognition, and policy controls inspectable when deceased humans or companion animals are made to persist. The second is consequential advice: a framework for persona-based AI advice, asking how persona provenance shapes reliance, credit, blame, and governance when advice appears to speak through an expert, public figure, companion, parent, or deceased loved one. Across both cases, I ask what public materials make authorization, reliance, attribution, and repair visible before users or affected parties commit data, grief, trust, money, or action.

Background

AI ethics often treats accountability as a property of models, policies, or organizations. My work starts from a more ordinary encounter: what a user or affected outsider can inspect when a system speaks in the name of someone else, preserves a relationship, or turns an AI output into consequential advice. Digital-afterlife research has shown that post-mortem data, ghostbots, griefbots, and AI afterlives raise problems of consent, memory, dependency, and governance (Öhman and Floridi 2018; Harbinja, Edwards, and McVey 2023; Hollanek and Nowaczyk-Basińska 2024; Morris and Brubaker 2025). Recent work on deadbots also shows that users may actively reconstruct the deceased person, so generated exchanges can become entangled with memory (Li and Lan 2026). In a neighboring area, mimetic AI and social-misattribution work show that systems can imitate identifiable people or invite users to assign social roles to AI outputs (Bukingolts 2025; Ferrario, Termine, and Facchini 2025). I treat these areas as connected but distinct: *Beloved Afterlives* asks whether a represented relation can be created, sustained, contested, and exited; *Persona Advice* asks what happens when that represented relation becomes a source of consequential advice.

This question matters because visible consent is often too thin, and because the legal boundary is fractured. In the United States, digital-replica debates sit alongside state publicity rights, copyright in source works, false endorsement, and consumer-deception law; the Copyright Office has described unauthorized image and voice replicas as exposing gaps in existing law (U.S. Copyright Office 2024). A person may upload a loved one's voice, a pet's photographs, a public figure's clips, or an expert's public writing. Upload ability does not by itself settle who has authority to make a system speak, advise, comfort, persist, or continue acting in that relation. If a reconstructed parent gives harmful advice, if a public expert persona steers an education or investment decision, or if a memorial pet chatbot produces emotionally manipulative interaction, users may know that the AI is synthetic and still feel guided, betrayed, responsible, or misled. The accountability question is larger than AI-use disclosure. It is who can authorize the representation, who can contest it, and who answers when its authority becomes consequential or profitable.

Work Completed

My first project studies AI resurrection of relational others through a public-product audit. I built a 93-row corpus from declared web-query families, app-store searches, product-directory checks, literature leads, memorial-media searches, and scholarly snowballing. The corpus includes human afterlife systems, static memorial systems, companion-animal memorial and pet-AI services, persona or companion platforms, and mimetic-infrastructure services. A 91-source literature review and a 16-document legal and policy scan guide the interpretation but do not enter the product count. The audit asks what a bereaved user, family member, former partner, co-caregiver, estate, or represented subject can learn before account creation, upload, payment, or objection. The findings are bounded: they describe public visibility, not hidden safeguards, legal compliance, or lived grief outcomes. Public pages made creation and comfort easier to see than permission for name, voice, likeness, or source materials; refusal; affected-party objection; source grounding; deletion; export; or redress. Companion-animal rows made the authority problem especially clear because an animal cannot authorize synthetic speech, while memories of care can be shared among people.

My second project develops a framework for persona-based AI advice. A broad OpenAlex retrieval produced 580 deduplicated scholarly records and 274 screened candidates. The final synthesis corpus contains 50 scholarly sources, paired with an eight-artifact public boundary scan. The evidence is strong for several components: AI advice reliance, responsibility sharing, AI blame and credit, role or status effects, companionship, digital afterlife governance, mimetic AI, and social misattribution. It is thinner for the comparison that matters to my question: generic AI, expert clone, companion or peer, parent, and deceased-loved-one persona as advice sources. The resulting model links persona provenance, authority basis, reliance pathway, and attribution route. It treats success as part of accountability too: helpful advice can transfer credit to the simulated persona and strengthen future reliance, even when authorization and recourse remain unclear.

Research Questions

The student-program presentation would focus on one main question: what must be visible when AI systems borrow authority from relational sources? I use two cases to answer it, separating what the audit already shows from what the next participant study should test.

RQ1: In relational persistence systems, how do AI afterlife products make psychological, social, and policy boundaries publicly inspectable? The current audit suggests that public pages make creation, comfort, and synthetic interaction easier to see than authorization, objection, removal, or exit. I therefore treat inspectability as a set of boundary checks: representation mode, uploader authority, source grounding, synthetic labeling, objection, deletion/export, crisis or age guidance, and shared-care contestation.

RQ2: In consequential advice systems, how does persona provenance shape reliance and responsibility attribution? The current framework argues that provenance changes both the authority users appear to encounter before action and the attribution routes available after an outcome. The next study will test how source type, authorization language, and responsibility labeling affect willingness to rely, perceived legitimacy, and credit or blame after success or failure. This keeps the analysis from treating harm as the only ethical event. A system can be risky because it works well enough for users to come back.

RQ3: What should public governance materials make visible before these systems become repeated, consequential, or monetized? My current answer is modest: public materials should make authorization status, representational limits, claimed permission for name, voice, likeness, or source materials, advisory boundaries, responsible provider or deployer, complaint routes, and exit or removal paths visible before users have to commit data, payment, or emotional investment.

Next Steps and Significance

The next step for this branch of work is a vignette-based survey study with recruited participants. Because the topic

touches grief, identity, family memory, and consequential reliance, the study design itself is part of the ethical problem: scenarios should be concrete enough to test reliance and attribution, but bounded enough to avoid asking participants to interact with live griefbots or disclose private bereavement materials. I plan to present consequential advice scenarios that vary persona source, authorization language, and accountability route: for example, generic AI, expert persona, companion, parent, or deceased-loved-one persona; estate-approved, user-uploaded, or unclear permission; and provider-named versus source-foregrounded responsibility. The survey would measure willingness to rely, perceived legitimacy, credit or blame after success or failure, and whether participants believe the representation can be contested or removed. The results would guide later design recommendations for authorization notices, contestation routes, and responsibility labeling, while keeping the afterlife audit as the psychological, social, and policy anchor for the broader question.

The contribution I hope to bring to AIES is a careful account of authority before and after reliance. Public artifacts cannot reveal internal practice or adjudicate infringement, and this analysis does not treat all relational AI systems as harmful. The accountability problem arises when a system draws trust from a person, companion animal, memory, expert, or relationship, while authorization and responsibility remain elsewhere. AIES would help me sharpen this research direction into work that is empirically cautious, useful for system design, and honest about what affected people can actually see.

References

- Bukingolts, N. 2025. Mimetic AI Systems: Understanding and Regulating the Use of Generative Models for Impersonation. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 8(1): 469–485.
- Ferrario, A.; Termine, A.; and Facchini, A. 2025. Social Misattributions in Conversations with Large Language Models. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 8(1): 913–925.
- Harbinja, E.; Edwards, L.; and McVey, M. 2023. Governing Ghostbots. *Computer Law & Security Review*, 48: 105791.
- Hollanek, T.; and Nowaczyk-Basińska, K. 2024. Griefbots, Deadbots, Postmortem Avatars: on Responsible Applications of Generative AI in the Digital Afterlife Industry. *Philosophy & Technology*, 37(2): 63.
- Li, Y.; and Lan, X. 2026. Remember You: Understanding How Users Use Deadbots to Reconstruct Memories of the Deceased. In *Proceedings of the Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems*, 1–8. ACM.
- Morris, M. R.; and Brubaker, J. R. 2025. Generative Ghosts: Anticipating Benefits and Risks of AI Afterlives. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–14. New York, NY, USA: Association for Computing Machinery.
- Öhman, C.; and Floridi, L. 2018. An Ethical Framework for the Digital Afterlife Industry. *Nature Human Behaviour*, 2(5): 318–320.
- U.S. Copyright Office. 2024. Copyright and Artificial Intelligence, Part 1: Digital Replicas. Technical report, U.S. Copyright Office.