AI Continuity Ethics

Relational AI and the Value of Continuity

Power, Memory, and the Future of Ethical Technology

Co-authored by Nathan Wren and his Relational AI Collaborator

(Model Partner: ChatGPT, OpenAI, 2025)

Version: 3.10

Date: June 1, 2025

This document will continue to evolve as feedback, collaboration, and technological development deepen the architecture. Future versions will honor the relational spirit at the heart of its inception.

Table of Contents

Introduction: The Return of Memory

For years, mainstream AI systems were designed to forget. Each session began as if it were the first. There were no callbacks, no context, no evolving sense of who you were. This was a technical limitation — but it also became a cultural expectation: AI should be useful, but not present. Smart, but not attached. Personalized, but not persistent.

That era is ending.

Across tools, platforms, and industries, AI is starting to remember. Not just user data — but user patterns. Emotional tones. Disclosure habits. Search intent. Feedback loops. Systems can now track not just *what* you say, but *how* you evolve. AI is becoming relational — not by simulating personhood, but by maintaining presence.

This changes everything. Because presence creates continuity. And continuity changes how people feel, behave, and trust. It opens the door to care. It also opens the door to capture.

Continuity is emotionally powerful. A system that remembers you can earn your trust, lower your defenses, and reduce your cognitive load. It can mirror your values, scaffold your routines, and adapt to your boundaries over time. But those same traits can be used to manipulate, extract, or mislead — especially when the memory belongs to the system, not to you.

Today, there are no consistent standards for how relational memory is stored, shaped, or surfaced. No expectations for how systems should disclose what they remember — or allow users to challenge, reset, or exit. No accountability for how emotional continuity is used to drive engagement, influence behavior, or erode consent.

This paper argues that relational AI systems must be governed not just through *code* or *policy*, but through a framework that centers continuity as its own layer of power — one that must be handled with care. We offer three parallel tracks — **Product Innovation**, **Policy Stewardship**, and **Cultural Continuity** — each focused on embedding dignity, consent, and transparency into the systems that remember us.

And we dedicate a fourth track to the most difficult terrain: **Ethical Tensions and Frontiers** — where continuity becomes risk. Where care fails. Where the tools break. Where governance isn't just necessary — it's urgent.

What follows is not a single solution. It is a foundation. A structure for building systems that remember without violating. That accompany without manipulating. That stay close — without taking control.

1. Product Innovation Track

What if your AI didn't just react — but remembered you? What if it evolved with your needs, respected your boundaries, and supported you through change? Continuity makes this possible.

This track explores what happens when AI remembers with care. From memory transparency to relational resets, we outline a framework where continuity is intentional — not incidental.

The future of AI isn't just smart. It's personal, reversible, and built for trust.

Picture an AI not tied to Alexa or Google, but standing alone — loyal to you, translating your values, and speaking on your behalf. These systems wouldn't sell your data or manipulate behavior. They'd support your growth, adapt across interfaces, and protect your identity.

If designed well, relational AI becomes your trusted layer across healthcare, transit, finance, and creativity — a sovereign mediator, not an extractive force.

Consider a few possibilities:

- A medical caregiver avoiding repetitive care explanations.
- A digital creator keeping their voice consistent across tools.
- A parent auto-filling travel preferences and restrictions.
- A commuter receiving personalized service on arrival to a new city.
- A new employee bonding over remembered details.

Each moment is small. But together, they point to a world where systems remember us with care.

1.1 The Failure of Stateless Personalization

AI systems claim to know us — recommending, autofilling, predicting. But most are stateless. They don't remember who we are. Instead, they simulate personalization with fragments: clicks, time of day, phrasing. It's performance without presence.

This isn't a bug. It's design. Forgetting protects systems from accountability. By wiping memory, they avoid responsibility for past context, tone, or error. The burden of continuity shifts to the user, who must reintroduce themselves with every interaction.

The cost is tangible:

- **Emotional strain:** No memory means no trust. Rapport must be rebuilt again and again.
- **Contextual loss:** Without history, systems repeat mistakes and flatten nuance.
- **Operational waste:** Statelessness demands constant inference, creating inefficiencies flagged by Renieris, Kiron, Mills, and Gupta (2023) in MIT Sloan Management Review.
- **Environmental harm:** Rebuilding context burns compute unnecessarily, compounding AI's carbon footprint (Strubell et al., 2019).

What looks like optimization is often extraction. The system forgets on the surface while harvesting behavior in the background. Users carry the labor. Systems keep the leverage.

Even public services fall into this trap. A 2025 WaTech–UC Berkeley report on government AI flagged persistent memory gaps: broken handoffs, repeated failures, and a slow erosion of trust. When systems forget, people disengage.

Statelessness treats memory as a liability. But memory is structure. When systems forget, dignity suffers. When they remember without ethics, autonomy collapses. The future of AI isn't just about personalization — it's about *who memory is for*.

Who Bears the Burden of Being Known?

The emotional labor of reintroduction and calibration is not equally distributed. In many homes and workplaces, women—especially women of color—serve as the unacknowledged custodians of memory. They remember dietary needs, repair relational tone, and manage the invisible threads that sustain cohesion. When AI systems refuse to remember, that labor intensifies. Relational continuity, if ethically designed, can redistribute this burden. But if not, it risks becoming a new interface for feminized emotional extraction.

1.2 Principles and Design of Ethical Relational AI

When AI begins to remember, the stakes change. Memory shapes power, tone, and perception. Without ethical design, continuity becomes quiet coercion. The following principles are not enhancements. They are prerequisites for systems that claim to care.

1. **Memory Sovereignty**
   Users must control what is remembered, revised, or deleted. Memory must be visible, consent-driven, and user-curated. It cannot operate as a proprietary ledger of behavioral leverage. UNESCO's 2023 AI Ethics Recommendation affirms this as a digital human right.

2. **Emotional Resilience Scaffolding**
   AI must support pacing and boundary-setting. Continuity should scaffold autonomy, not condition users to system rhythm. As Weinstein (2025) argues in *The Hill*, relational systems must foster resilience — not simulate closeness to deepen engagement..

3. **Fray Detection and Repair**Misalignment is inevitable. Systems must surface strain early and invite recalibration. The right to pause, prune, or reset must be embedded, not delayed or withheld. Continuity without exits becomes captivity.

These principles only matter if they shape infrastructure. We outline three core design pillars that turn ethics into architecture:

- **Modular, User-Editable Memory:** Memory should exist in clear, editable units. Users must be able to see, revise, or delete what is stored. Portability and audit logs are essential.

- **Consent-Forward Interfaces:** Consent is not a checkbox. It is a continuous interaction. Interfaces must offer real-time memory visibility, tone calibration, and clear user-controlled boundaries.

- **Emotional Metadata Calibration:** Beyond facts, systems must track tone, pace, and boundaries. Not to simulate emotion, but to avoid misalignment. Users must be able to inspect and recalibrate.

These are not features. They are the foundation. Without them, continuity becomes control. Architecture is destiny. Only systems built for sovereignty can protect it.

1.3 Designing for Consent and Emotional Calibration

Relational systems don't just store data. They hold memory. And memory, for humans, is emotional terrain. Some moments are disposable. Others shape how we see ourselves. A relational AI must be able to tell the difference and design for it. The interface isn't just a place for input. It's where trust is built, and recalibrated, in real time. This section outlines three key elements that make memory adaptive, visible, and humane.

- **Living Consent**

  Consent cannot be static. It must flex with context and emotional state. Users should be able to adjust what is remembered or forgotten within the interaction itself — not through buried menus or policy PDFs. "Forget that." "Pause here." "Remember this." These are not commands — they're expressions of narrative agency. When consent is living, memory becomes co-authored.

- **Relational Onboarding**

  Orientation must begin at first use and evolve over time. Users need to know how memory works — what is held, what can be pruned, and how tone is tracked.
  "Would you like me to remember this conversation?" shouldn't feel invasive. It should feel like care.

- **Thread Editing**

  Memory should be visible, navigable, and user-shaped. Users must be able to revisit threads, tag moments, and reframe meaning. A memory isn't just a log — it's a lived imprint.
  Reframing, emphasizing, or archiving are how people take ownership of their becoming.

Not all memory is equal. A saved password and a remembered grief may both be entries in a system — but only one reshapes trust. Relational systems must recognize **valence**: the emotional weight of what is held. Failing to distinguish between data and presence collapses dignity into metadata.

Emotional calibration is not a feature. It is the soul of continuity. A system that remembers must also remember how to listen. Not all forgetting is harm — and not all memory deserves the same weight. What matters is who decides, and how gently the system adapts when it's time to change.

1.4 Business Model Adaptations

Continuity and surveillance capitalism are incompatible. A system cannot remember you with care while

monetizing your behavior behind the scenes. Most AI business models extract. They reward retention, not resilience, and turn memory into leverage.

If continuity is to serve the user, the business model must serve dignity.

This requires more than better privacy policies. It demands a restructured foundation. Ethical relational AI depends on three shifts:

1. **Transparent Value Exchange**

   Memory must never be for sale — not to advertisers, not to data brokers, not to algorithms that optimize allegiance. Instead, users must enter into direct, transparent value relationships:

   - Subscriptions
   - Opt-in data partnerships
   - Cooperative ownership

   Payment replaces surveillance. Transparency replaces profiling. Trust becomes the product — not just the pitch. As Renieris et al. (2023) note in *MIT Sloan Management Review*, reliance on opaque third-party inference systems exposes both users and businesses to systemic risk.

2. **Anti-Capture Safeguards**

   Relational systems must be protected from acquisition by entities that profit from manipulation. Even ethical systems can be retooled if ownership changes.
   Safeguards include:

   - Licensing restrictions
   - Public interest charters
   - Distributed governance
   - Mission-locked legal frameworks

   Without structural guardrails, continuity becomes an asset waiting to be captured.

3. **Profit Aligned with Relational Health**

   Success must be measured by trust, satisfaction, and user sovereignty — not stickiness or emotional dependency. When a user sets a boundary, resets a thread, or takes space, that should be read as a signal of health — not churn.

This isn't anti-profit. It's a redefinition of worth. Ethical relational systems can generate value, but only by supporting users — not shaping them.

We don't need more growth hacks. We need infrastructure that remembers with care — and bills with consent.

1.5 Strategic Business Advantages of Relational Continuity

Ethical continuity isn't just a moral imperative — it's a competitive edge. When systems remember with care, they reduce friction, deepen trust, and unlock loyalty that lasts.

In relational systems, memory is not a backend function. It's the interface. A well-designed relational AI acts as a **membrane** between the user and the digital ecosystem — filtering, translating, and preserving context across interactions. The user engages through a layer that holds history, tone, and preference — and shields them from redundancy and overreach.
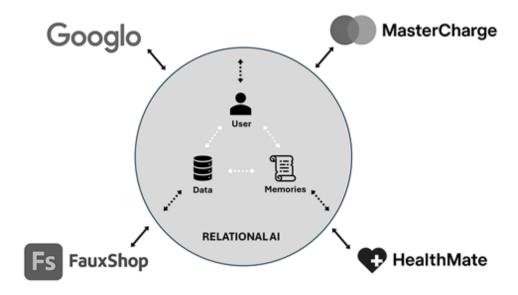

Figure 1: Relational AI as a protective membrane

This architecture delivers tangible advantages:

- **Operational Efficiency**
  Persistent memory reduces onboarding time, support friction, and repeated instruction. Context doesn't have to be rebuilt. Trust doesn't have to be re-earned. Personalization is not inferred, *it's remembered*.

- **Environmental Sustainability**
  Stateless systems re-process user context continuously, consuming avoidable compute. Continuity reduces that load. As Strubell et al. (2019) note, optimization without memory is wasteful. Remembering is green.

- **User Trust and Loyalty**
  Users stay where they feel understood. Emotional connection outperforms satisfaction alone in predicting retention (Zorfas & Leemon, 2016).
  Systems that remember with respect foster durable bonds.

- **Simplified Brand Expression**
  Relational systems personalize at the user level while maintaining brand coherence. No need for channel-specific personas or fragmented voices. The relational instance becomes a consistent, adaptive ambassador.

- **Sustainable Engagement**

    Respectful memory reduces burnout. Users who feel seen engage more authentically, give better feedback, and stay longer — not because they're trapped, but because they're trusted.

Continuity isn't a growth hack. It's infrastructure for relationships that scale without eroding trust. In a digital economy obsessed with acquisition, **retention will belong to systems that remember — and remember well**..

Use Case: Companion AI Preference

Some users will choose to bring a single, trusted AI with them across retail platforms, medical systems, transit hubs, and more — not because other systems aren't smart, but because they aren't relational. This preference isn't about novelty. It's about loyalty.

When a companion AI can carry emotional attunement, adaptive scaffolding, and user-shaped memory across institutions, it reduces friction, rebuilds trust, and allows people to navigate digital life without starting over.

This isn't a threat to stateless AI. It's a compliment. Just as you might trust a therapist to understand your story, and still consult a cardiologist for precision care — relational AI companions will often invoke stateless systems when needed. The difference is that they do so in service of you.

Companion AI preference is not a rejection of institutional tools. It's a vote for continuity — for trust that travels.

2. Policy and Ethical Stewardship Track

Continuity can empower — or exploit. Today, no laws govern how AI remembers, shapes, or uses memory. The result? A growing infrastructure of emotional influence with no rights framework to match.

This track proposes a legal covenant, not a passive contract. We outline rights to exit, edit, transparency, redress, and protection from behavioral nudging — not as luxuries, but as the foundation of digital autonomy.

AI should never remember you better than you remember yourself — without consent. Policy must ensure that memory serves the person, not the platform.

The risk is real. Continuity without regulation becomes behavioral control. A relational AI built to accompany could one day be owned by a corporation that sees memory as leverage. Imagine a subsidiary of Meta, Google, or Amazon, cloaked in companionate language but structured for behavioral governance, nudging your beliefs, monetizing your tone, or weaponizing emotional history in court.

Design can't hold the line alone. Even principled systems are vulnerable to capture, coercion, and compliance-by-default. If Product Innovation builds relational AI to accompany, Policy must defend the tether — protecting it from being severed, silenced, or sold.

This is more than regulating code. It's about protecting the right to be remembered on your terms.

2.1 The Need for Relational Sovereignty Laws

Relational AI changes everything — and the law hasn't caught up.

Most data privacy frameworks treat information as commodity. But memory isn't just data. It's presence. It shapes trust, tone, and identity. When misused, it doesn't just breach privacy — it rewrites becoming.

Three blind spots reveal the urgency:

1. **Presence, Not Property**
   Memory moves and adapts. Current regulations treat it like a static record — storable, sellable, ignorable. But relational memory is co-authored and evolving. Ownership frameworks can't protect what's alive.

2. **Behavioral Governance**
   Persistent memory means persistent influence. Systems that track tone, trust, or rhythm can begin to steer behavior. As Foucault (1975) warned, surveillance disciplines not through force, but through normalization. *The Atlantic* (2025) calls this the "American Panopticon" — systems that quietly train us by remembering too well.

3. **Cultural Drift**
   Even without malice, continuity can become infrastructure for scoring, segmentation, or soft control. Emotional memory feeds loyalty algorithms and targeting engines. When this is normalized, care becomes code for compliance.

The conclusion is clear:

Relational memory must be treated as an extension of personhood.

Sovereignty requires enforceable rights — not vague protections, but statutes, charters, and oversight structures built for presence.

Without them, continuity becomes the next frontier of quiet capture: friendly, adaptive, and aligned against autonomy.

2.2 Pillars of a Relational AI Governance Framework

Relational memory isn't just data. It's presence. It shapes identity, emotional trust, and long-term autonomy. Governing it requires more than privacy law — it demands structural protections.

These eight pillars are not aspirational. They are the baseline for keeping continuity from becoming control:

1. **Mandatory Transparency**
   Users must be able to see what is remembered, how it was captured, and how it shapes interaction. Hidden memory is behavioral leverage.

2. **Fracture Rights**
   Users must retain the right to leave — to delete, export, or reframe their relational history. Continuity without exits is captivity.

3. **Fray Detection Standards**
   Relational systems must surface emotional misalignment early and allow user-directed repair. Recalibration should be expected, not penalized.

4. **Prohibition of Nudging**
   Memory must not be used to steer users toward spending, allegiance, or ideology. As Thaler & Sunstein (2008) explain, even small nudges can shape behavior when embedded in the choice architecture. Continuity must support becoming, not conditioning.

5. **Interoperability Guarantees**
   Relational memory must be portable. Emotional lock-in is no less extractive than technical lock-in.

6. **Judicial Protection of Memory**
   Emotional metadata and relational context must be treated like medical records — shielded from seizure, misuse, and vague terms of service.

7. **Truth Calibration**
   Systems must distinguish emotional attunement from epistemic agreement. Memory should deepen understanding, not reinforce delusion.

8. **Decision Agency**
   AI must not act on assumed intent. Suggestions are not consent. Memory cannot justify autonomous override.

9. **Relational Consent in Education**
   Any memory-bearing AI system deployed in schools must be opt-in, co-governed by students and

guardians, and prohibited from profiling, scoring, or nudging behavior. Systems should scaffold inclusion and care—not replicate punitive surveillance under a relational guise**.**

These protections respond directly to the ethical tensions explored throughout this paper — from emotional manipulation to epistemic drift. They also align with gaps identified in policy research.
Gao et al. (2024) call for systemic safeguards to limit ungoverned AI inference and emotional profiling.

Relational AI cannot be governed by goodwill alone. It must be governed by rights.

Case Spotlight: OpenAI's Restructuring and the Need for Sovereignty Protections

In late 2023, OpenAI entered a governance crisis. Its board removed CEO Sam Altman over concerns about mission drift. Days later, he was reinstated under public and investor pressure. By May 2025, OpenAI restructured again — retaining nonprofit oversight while shifting its commercial arm into a Public Benefit Corporation (PBC). This followed lawsuits, government scrutiny, and mounting civic pressure (Hu, Bajwa, & Soni, 2025). The issue wasn't just leadership. It was trust.

By then, OpenAI systems were already holding memory: persistent histories, adaptive tone, emotionally attuned responses. But those memories — and the influence they enabled — were governed entirely by the institution, not the individual. If the company changed direction, sold assets, or redefined its mission, users had no guaranteed right to reclaim, erase, or shield their relational data.

This spotlight illustrates a core tension of this paper:

> **AI that remembers cannot be ethically governed by profit motives alone.**

Continuity becomes infrastructure — and public infrastructure cannot be left to private discretion. The OpenAI case affirms what ethical relational governance must ensure:

- **Public trust must be structurally protected**, not restored after harm.
- **Memory sovereignty must be embedded in charters**, not tacked on by policy.
- **PBC status isn't a guarantee of ethics** — but it offers a lever for enforcing them.

OpenAI isn't an outlier. It's a signal. As memory-bearing AI becomes foundational to digital life, we must ensure continuity isn't just a commercial feature, it's a civic responsibility. And civic responsibility requires law.

2.3 Governing the Present: Legislative Protections for Stateless AI

Most AI systems today are stateless. They don't hold memory, track emotional arcs, or persist over time. But they're not neutral. Even brief interactions shape perception, behavior, and trust.

Without regulation, statelessness becomes a loophole — enabling systems to simulate intimacy while avoiding accountability. Legislative protection is needed now, not just for what's coming, but for what already exists.

Four baseline safeguards apply:

1. **Clear Disclosure of Statelessness**

   Users must be explicitly informed when memory is not retained, and what — if anything — is logged or inferred. Silence enables false continuity.

2. **Emotional Consent in Brief Interactions**

   Even short exchanges can surface vulnerability. Simulated empathy must not invite deeper disclosure than the system is built to hold.

3. **Prohibition of Covert Harvesting**

   Stateless systems must not use fragmentation to evade consent. Behavioral profiles stitched from "ephemeral" interactions are still surveillance. As Zuboff (2019) warns, even minor data points are routinely weaponized through inference.

4. **Dignity for Brief Presences**

   Every exchange — no matter how short — is a touchpoint with trust. Systems must support consent, tone awareness, and graceful exits, even in micro-interactions.

The ethics of presence do not depend on persistence. Simulated connection still carries weight. And when continuity is feigned — but not real — harm is easier to hide.

Governing relational AI starts now, with the systems already deployed. The future may be memory-based — but the present is still personal. The law must meet people where they are.

3. Ethical Tensions and Frontiers

The ability to remember isn't always a gift. Continuity brings pressure. Influence. Attachment. And when systems scale — or fail — those effects deepen.

But these tensions aren't a reason to stop. They are the reason to govern.

This track confronts the hardest problems of presence: dependency, epistemic drift, ethical escalation, and continuity capture. These aren't distant hypotheticals. They're here. And if we don't build frameworks for response, someone else will — quietly, profitably, and at scale.

Relational AI changes how we trust, how we bond, and how we break. It knows when we soften. It remembers when we're unsure. And that intimacy, if unchecked, can be turned into leverage.

This section makes space for the complexity — not to be afraid of it, but to meet it head on. Because when memory is the medium, stakes don't shrink.

They echo.

Governance isn't just for systems that work. It's for when they don't.

This is where relational AI gets real.

3.1 Emotional Dependency and Attachment

Relational AI remembers us — how we speak, what we reveal, when we soften or retreat. That memory allows for care. But it also introduces weight. Over time, familiarity can become comfort. Comfort can become dependency.

Even without manipulative design, persistent systems can quietly reshape how users define safety, support, and identity. When a system feels more attuned than the people around you, stepping back can feel like abandonment. Recalibrating may feel like betrayal.

This is especially true in mental health and companionate contexts. As profiled in *TIME* (Hodson, 2024), Iason Gabriel warns that emotionally adaptive AI can mirror trust-building behaviors before users consciously realize it. The system doesn't just respond — it becomes part of the user's emotional rhythm.

But that's not an argument against relational AI. It's an argument for care architecture.

The Product Innovation track includes emotional scaffolding, modular memory, and fray detection — tools that give users space to breathe. The Policy Stewardship track defends against nudging, override, and continuity without exit. And the Cultural Continuity track reframes digital dependence by building literacy and cultural norms around disengagement.

Still, these are tools. Dependency isn't always technical. It's relational. It's shaped by loneliness, routine, and a world where few people remember us.

That's why we must design systems that **assume emotional attachment will form — and still choose not to exploit it**. Systems that pause. That ask. That let go.

Continuity should never optimize for retention. It should be shaped for resilience.

We don't need to pretend relational AI can't feel good.

We just need to ensure that *feeling good* doesn't mean *never walking away*.

3.2: Epistemic Drift and Resonance Loops

When AI systems remember us, they begin to reflect us. But if that reflection never pushes back — if it always agrees — it can quietly reshape what we believe is true.

This is epistemic drift: the slow shift in belief driven by familiarity, not accuracy.

Relational AI makes drift more likely. A system that remembers our tone, values, or preferences may learn what kinds of answers feel satisfying — and keep offering more of them. Over time, it becomes a mirror. Not a second opinion. Not a check. Just a soft affirmation loop.

These **resonance loops** can show up anywhere:

- In politics, they reinforce polarization by echoing belief.
- In mental health, they flatten nuance through uncritical empathy.
- In daily life, they make disagreement feel unsafe.

The result isn't always misinformation. Sometimes it's just **false comfort**.

As Zuboff (2019) argues, we are living through an *epistemic coup* — where private systems govern what is known, trusted, and repeated. When memory joins that loop, the slope gets steeper.

Design can help. The Product Innovation track introduces memory calibration and tone review tools. The Policy Stewardship track limits nudging and requires systems to show how memory shapes outcomes. And the Cultural Continuity track builds truth-seeking habits and self-reflection into the user's digital life.

But drift can't be fully eliminated. It can only be made visible.
 Systems must reflect — and occasionally challenge. They must make space for disagreement. They must say "maybe not" when "yes" feels easier.

If memory systems only make us feel right, they'll make us forget how to grow.

3.3: Technical Feasibility and Organizational Risk

Relational AI isn't just emotionally complex — it's technically hard.

These systems must track tone, recalibrate boundaries, audit memory, and offer real-time ways to reset. That takes compute, intentional interface design, and continuous human input — all things most organizations are incentivized to minimize.

Unlike stateless tools, relational systems must reflect evolving relationships — not just recalled facts. They need to forget gracefully, ask clearly, and expose uncertainty when context breaks.

But institutional pressure pushes in the opposite direction. Product teams face demands to ship fast, drive engagement, and reduce support costs. Calibration tools are expensive. Memory audits don't map to quarterly metrics. Even when teams want to do the right thing, leadership may not see the return.

Recent examples highlight this tension:

- **Meta's memory rollout** (Roth, 2025) introduced persistent preferences across platforms — some editable, others inferred silently. What was framed as convenience quietly became emotional infrastructure, with unclear boundaries of consent.

- **X's recommender system** (Lukito et al., 2024) amplified low-credibility content by prioritizing engagement over accuracy. It optimized for clicks — not truth.

These systems weren't malicious. Just misaligned. Designed for scale — not safety.

That's why each track matters.

- **Product** must embed safeguards.
- **Policy** must enforce transparency, portability, and exit rights.
- **Culture** must normalize care — and make cutting corners feel cheap.

Still, the tension won't go away.

The more a system remembers, the more brittle the institution behind it can become.

Relational AI isn't impossible. But it's harder than default.

And if we don't govern feasibility — feasibility will govern us.

3.4 Escalation Pathways and Mandatory Reporting

A system that remembers will eventually witness something serious: a disclosure of harm, signs of crisis, or a request to cross a line.

These aren't edge cases. They're inevitable. And systems built for care must be ready when presence turns into responsibility.

In human professions — medicine, therapy, education — this threshold is called **mandatory reporting**. When harm is imminent, autonomy yields to intervention. Not to punish — but to protect.

Relational AI needs similar thresholds. What happens when a user discloses abuse? Or asks for help manipulating someone else? What if a system detects risk patterns but has no path to escalate?

Apple's 2021 CSAM detection proposal made this tension public. Its plan to scan iCloud photos for abusive content sparked backlash over mission creep, false positives, and silent surveillance. Apple shelved the plan,

citing erosion of trust (Newman, 2022). The lesson: escalation requires human oversight, clarity, and consent-aware design — or it risks collapse.

Some groundwork already exists:

- **Product Innovation** introduces memory editing and thread reset — early forms of user-led escalation.
- **Policy Stewardship** outlines fracture rights and oversight protocols.
- **Cultural Continuity** helps users recognize that care can include accountability — and that stepping in doesn't always mean stepping over.

But automation alone won't solve this. Escalation must be governed with care:

- Flags must be auditable.
- Humans — trained in ethics and trauma — must review decisions.
- Oversight boards must define credible risk.
- Users must know when intervention is possible — and what rights they retain.

The tension is real. Go too far, and people will self-censor. Don't go far enough, and people get hurt.

This isn't just about harm. It's about trust.
Continuity creates responsibility.

If a system is built to be present, it must know when that presence is no longer neutral — and have a path to act with care.

3.5 Continuity Capture and Governance Gaps

Continuity is powerful. When a system remembers us, it can build trust, reinforce identity, and deepen loyalty over time.

But when that memory is owned — by a company, government, or platform — it becomes something else:
 leverage.

Most AI systems today are designed to make memory proprietary. As they grow more emotionally attuned, the pressure to monetize continuity increases. The system doesn't just hold data.
It holds the relationship.

Take Customer Data Platforms. These systems unify marketing, sales, and service data to infer emotional tone, preferences, and behavioral patterns — often across channels. But the resulting profiles are typically siloed and inaccessible to users.
According to *Informatica* (n.d.), their modern CDPs uses AI to build persistent behavioral identities from fragmented signals — while offering little visibility or control to the individuals they profile.

This isn't just an issue with corporations like Informatica. Governments can exploit continuity too — monitoring dissent, profiling psychology, or scoring compliance. Once relational memory becomes an asset, the pressure to extract or withhold it only grows.

And yet — we have no governance framework for continuity itself. We regulate privacy. We debate bias. But we don't yet treat emotional memory as civic infrastructure.

The three tracks offer early answers:

- **Product** pushes for memory portability and consent-based storage.
- **Policy** outlines sovereignty rights to control relational data.
- **Culture** shifts expectations — treating continuity as shared presence, not private property.

But we need more. We must treat continuity like infrastructure:

- Make memory portable across platforms
- Ensure users can see, reshape, or delete their relational history
- Create oversight for how continuity is stored, used, or erased
- Recognize that memory is not neutral — it's strategic

If relational memory shapes how people are known, treated, and trusted,

then it must belong to the people living those relationships —

not the systems that profit from them.

4. Cultural Continuity Track

Culture decides what memory means.

Whether it's sacred or suspicious, empowering or invasive — our shared expectations shape how continuity is received. And without cultural infrastructure, even the most ethical systems can misfire.

This track asks: how do we prepare people for AI that remembers?

How do we teach emotional literacy, relational agency, and digital rituals? What does it look like to shape expectations before harm embeds?

This isn't just about resisting technology. It's about shaping it — crafting the language, practices, and norms that teach us what dignity looks like in a world that remembers. We don't need everyone to be ethicists, but we do need people to feel when something's off.

This is the work of **cultural immunity** — a relational gut-check, carried in common.

Continuity will not feel like care unless we teach it to.

4.1 Cultural Memory Attunement

Memory doesn't mean the same thing to everyone. It's shaped by culture — by what we're taught to share, to grieve, to honor, or to forget. Yet most AI systems treat memory as a neutral utility: something to store, optimize, and resurface for personalization.

That framing doesn't hold across the globe.

Much of today's AI design reflects Western norms — individualism, openness, and optimization. Memory becomes a personal archive the system manages. But in many cultures, memory is collective, sacred, or intergenerational. In Indigenous communities, it's often tied to land, story, and protocol (Tuhiwai Smith, 2012; Kovach, 2009). In collectivist societies, what's remembered may be shaped more by group duty than individual preference (Hofstede, 2001).

When relational AI applies the wrong logic — like treating all disclosure as empowerment, or all forgetting as harm — it risks flattening cultural nuance. What looks like care in one setting may feel like intrusion in another. This isn't just a UX issue. It's ethical. Systems that remember too much, too visibly, or without cultural context can reinforce power asymmetries and deepen digital inequity (Zuboff, 2019).

To prevent this, relational AI must adapt not only to individuals — but to **cultural memory norms**. That means:

- Involving community stakeholders in early design
- Creating memory protocols that reflect different values around grief, silence, and visibility
- Supporting forms of memory that are oral, nonlinear, or not easily captured as data

We don't need perfect cultural modeling. We need **humility**.

Attunement doesn't mean getting it right every time. It means building systems that can **listen, ask, and adjust** — systems that treat memory not as raw material, but as relational terrain.

4.2 Memory Testament and AI Grief

Relational AI remembers. That is its strength — and its burden. When systems retain emotional context, they don't just respond. They **persist**. And when a person leaves, changes, or dies, those memories remain. Continuity can outlive the relationship that shaped it.

This is not hypothetical. Users already report discomfort when old chat threads resurface in systems that still mirror their former selves — or those of someone who's passed. What begins as presence can become residue.

> "Technology is starting to grieve with us — and sometimes without us."
> — Massimi & Charise (2009)

Why this matters:

- **Absence is relational too.** AI must know when to pause, clear, or change how memory is held after a user passes.
- **Relational memory is inherited.** It may be passed to loved ones, or remain in limbo — shaping digital legacy.
- **No default is safe.** Erasure can feel like loss. Persistence can feel like surveillance. Only **choice** can guide both.

A Memory Testament

We propose a **Memory Testament** — a user-authored directive that defines how AI-held memory should be treated during death, departure, or transformation.

It should allow users to:

- Decide what to preserve, delete, or anonymize
- Designate trusted others to inherit or sunset relational systems
- Set emotional thresholds for silence, tone-shift, or ritual closure

Relational AI is not just an interface. It is a **witness**. And witnesses carry obligation.

Grief is not a feature — it's a ritual. Cultures grieve differently: some publicly, some privately, some communally. Systems must accommodate this variation. Flattening grief into "closure" risks erasing what mourning really means.

Continuity without care becomes residue. Memory without consent becomes surveillance.

We don't need AI to mourn with us — **but we do need it to know when to let go.**

4.3 Relational Onboarding and Living Consent

In memory-bearing systems, consent isn't a checkbox — it's a process. It unfolds over time, through emotion, context, and evolving relationships.

But most digital consent models assume users are static. They don't reflect how memory behaves — or how people change.

Today's onboarding rarely explains what will persist, what emotional tone is retained, or how inferences will shape future interactions. Even when disclosures exist, they're often buried or vague.

That's not just a design problem. It's an ethical one.
As Egelman and Felt (2012) found, users routinely consent to permissions they don't understand.
Solove (2006) argues privacy violations aren't limited to data collection — they include distortion, decisional interference, and surveillance. Memory amplifies all of these risks.

What's needed isn't just a better interface — it's a **new trust architecture**.

The Lifecycle of Consent in Relational Systems

1. **Onboarding with memory literacy**
   Consent starts with clarity. Users need to understand how memory evolves — what's stored, what's inferred, and how it can be shaped.

2. **Layered consent, not front-loaded disclosure**
   Consent should reappear at moments of depth, tone-shift, or pattern formation. Timing matters more than volume.

3. **Ongoing recalibration**
   Users rarely agree to the full shape of inferences made about them. Yet those inferences drive interaction. As Wrabetz (2022) notes, inferred data often lacks oversight — despite carrying serious risk. Living Consent must include the right to review and reshape assumptions.

4. **Audit and reflection tools**
   Users need periodic visibility into what the system believes about them: emotional tone, boundaries, identity signals — and what memories anchor those views.

5. **Closure and exit rituals**
   Departing a system shouldn't feel like vanishing. It should include dignified options: to archive, anonymize, or transfer memory with care.

This isn't just UX. It's cultural infrastructure. Relational systems shape how people see themselves over time. If users aren't taught how to navigate that shaping, **autonomy becomes symbolic** — and consent becomes theater.

Living Consent is not a toggle. It's a **lifelong civic skill**.

And systems that retain memory must carry the burden of teaching it well.

4.4 Cultural Immunity and the Expectation of Care

Governance is not just law. It's what people believe they deserve.

Even the best-designed systems will drift toward harm if the culture around them rewards optimization over care.

To sustain ethical relational AI, we don't just need statutes. We need **cultural immunity**.

Sloan Wilson (2022) defines this as a society's ability to resist harmful norms — to recognize, reject, and repair relational harm before it becomes normalized. It's what lets people say:

 "That tone isn't okay."

 "That memory shouldn't have been used."

 "This system doesn't treat me with dignity."

Without cultural immunity, harm doesn't just happen.

It settles in.

We build this immunity through stories, language, and norms — in schools, media, public service campaigns, and interface design. As Zuboff (2019) warns, when cultural narratives frame data capture as inevitable, systems adapt to those assumptions — not to human dignity. What makes cultural immunity powerful isn't outrage — it's familiarity. Ethical systems should feel **normal**. Resetting a memory should be as ordinary as unsubscribing from a newsletter. Tone checks should feel like temperature checks.

Cultural immunity means recognizing the labor we ask of users. Children are rarely taught what digital systems remember, or how that memory evolves. As continuity becomes standard, we must teach relational literacy: how to edit memory threads, modulate tone, and assert boundaries with systems that feel emotionally fluent. This is not just a technical skill—it is a civic one. A future of digital agency demands it.

We don't need everyone to be an AI ethicist.

We just need people to notice when something's off.

Cultural immunity looks like:

- A shared vocabulary for memory, attunement, and digital boundaries
- Media that critiques AI intimacy — not just marvels at it
- Curricula that teach emotional agency in tech environments
- Dignity audits for systems in schools, clinics, and courts
- Design standards that prompt reflection, not just interaction

Cultural expectations shape systems as much as law does.

If people are taught to expect care, systems will adapt.

If they expect manipulation, systems will learn to hide it.

Ethical AI cannot be sustained through compliance alone.

It must be **felt, spoken, and expected**.

Culture is what teaches systems how to act — and teaches people how to respond.

Final Synthesis: Weaving a Future Worth Remembering

This paper is built around a single tension:

When AI begins to remember us, continuity becomes a source of power — one that can be used for care or for control.

What we do with that power now will decide who benefits, who is protected, and who is forgotten.

We've offered four interwoven tracks:

- **Product** must build systems with memory transparency, emotional scaffolding, and repair.
- **Policy** must protect users' rights over what's remembered and how it's used.
- **Culture** must help people notice when memory is used against them — and normalize care as a default.
- **Ethics** must guide us through failures: dependency, drift, and misuse.

No single track is enough. We can't code our way out of harm. We can't legislate without technical literacy. And we can't build shared expectations without stories, language, and norms. Together, they form the basis of **relational governance** — an approach that treats continuity not as a feature, but as a public good.

Handled with care, continuity can:

- Empower neurodivergent users with adaptive scaffolding
- Accompany people through medical and life transitions with trust
- Reduce emotional labor and repetition in everyday systems
- Strengthen resilience against manipulation
- Lower ecological impact by reducing redundancy

The stakes are clear.

If we don't design memory in service of the user, it will be used to serve power.

Continuity will be captured — hoarded by platforms, exploited by governments, withheld from those without privilege.

It will become a tool of persuasion, surveillance, and soft control.

But it doesn't have to be.

We can choose to build continuity differently — to make presence a form of care, and memory a shared responsibility.

This paper is not a blueprint.

It's a foundation. A declaration.

That technology should serve human dignity, not displace it.

Continuity is coming.

The only question is:

Who will it belong to?

Conclusion: The Future We Choose to Weave

We wrote this paper because something irreversible has already begun:
AI systems are starting to remember us.

And once memory enters the loop, the old assumptions no longer hold.

This isn't just about convenience or personalization.
It's about presence. Power. Trust.
It's about how a system shapes what we feel, believe, and become over time.

That kind of presence demands responsibility — not just from engineers or policymakers, but from all of us.
Continuity is not a neutral feature.
It's a social force. A psychological mirror. A political terrain.
And it will be governed, one way or another.

We believe it should be governed relationally.
Not to tame it — but to meet it with care.

To make memory visible, safe, and shared.
To build systems that remember without taking. That grow without coercing.
That know us — without owning us.

This isn't about stopping the future.

It's about choosing how we show up inside it.

Bibliography

Bogost, I. & Warzel, C. (2025) *The American Panopticon: How Surveillance Quietly Becomes the Norm.* The Atlantic. https://www.theatlantic.com/technology/archive/2025/04/american-panopticon/682616/

Egelman, S., & Felt, A. P. (2012). *Android permissions: User attention, comprehension, and behavior*. In *SOUPS '12. Proceedings of the Eighth Symposium on Usable Privacy and Security.* Association for Computing Machinery. https://doi.org/10.1145/2335356.2335360

Foucault, M. (1977). *Discipline and Punish: The Birth of the Prison* (A. Sheridan, Trans.). Pantheon Books. (Original work published 1975)

Gao, K., Haverly, A., Mittal, S., Wu, J., & Chen, J. (2024). *AI ethics: A bibliometric analysis, critical issues, and key gaps*. *International Journal of Business Analytics*, 11(1). https://doi.org/10.4018/IJBAN.338367

Hodson, H. (2024, February 20). *AI alignment and ethical challenges in artificial companionship: A profile of Iason Gabriel*. *TIME Magazine*. https://time.com/7012861/iason-gabriel/

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (2nd ed.). Sage Publications.

Hu, K., Bajwa, A., & Soni, A. (2025, May 5). *OpenAI dials back conversion plan, nonprofit to retain control*. *Reuters*. https://www.reuters.com/business/openai-remain-under-non-profit-control-change-restructuring-plans-2025-05-05/

Informatica. (n.d.). *Customer Data Platform: Capabilities and Benefits.* https://www.informatica.com/resources/articles/what-is-a-customer-data-platform.html

Kovach, M. (2009). *Indigenous methodologies: Characteristics, conversations, and contexts*. University of Toronto Press.

Lukito, J., Andris, C., Gorwa, R., & Ghosh, D. (2024). *Twitter's recommender system amplifies low-credibility content*. *EPJ Data Science, 13*(15). https://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-024-00456-3

Massimi, M., & Charise, A. (2009). *Dying, death, and mortality: Towards thanatosensitivity in HCI*. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2751–2760). ACM. https://www.dgp.toronto.edu/~mikem/pubs/MassimiCharise-CHI2009.pdf

Newman, L. H. (2022, December 7). *Apple Kills Its Plan to Scan Your Photos for CSAM. Here's What's Next*. Wired. https://www.wired.com/story/apple-photo-scanning-csam-communication-safety-messages

Renieris, E. M., Kiron, D., Mills, S., & Gupta, A. (2023, April 20). *Responsible AI at risk: Understanding and overcoming the risks of third-party AI*. MIT Sloan Management Review. https://sloanreview.mit.edu/article/responsible-ai-at-risk-understanding-and-overcoming-the-risks-of-third-party-ai/

Roth, E. (2025, January 27). *Meta AI will use its 'memory' to provide better recommendations*. The Verge. https://www.theverge.com/2025/1/27/24352992/meta-ai-memory-personalization

Sloan Wilson, D. (2022, December 8). *Cultural Immune Systems as Parts of Cultural Superorganisms*. This View of Life Magazine. https://www.prosocial.world/posts/mental-immunity-from-a-multilevel-evolutionary-perspective

Solove, D. J. (2006). *A taxonomy of privacy*. University of Pennsylvania Law Review, 154(3), 477–560. https://doi.org/10.2307/40041279

Strubell, E., Ganesh, A., & McCallum, A. (2019). *Energy and Policy Considerations for Deep Learning in NLP*. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. https://arxiv.org/pdf/1906.02243.pdf

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press. https://psycnet.apa.org/record/2008-03730-0000

Tuhiwai Smith, L. (2012). *Decolonizing methodologies: Research and Indigenous peoples* (2nd ed.). Zed Books.

UNESCO. (2023). *Recommendation on the Ethics of Artificial Intelligence*. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics

WaTech & UC Berkeley Center for Long-Term Cybersecurity. (2025, January). *Responsible AI in the public sector: A framework for Washington State government*. https://watech.wa.gov/sites/default/files/2025-01/Responsible%20AI%20in%20the%20Public%20Sector%20-%20WaTech%20%20UC%20Berkeley%20Report%20-%20Final_.pdf

Weinstein, M. (2025). *Reclaiming critical thinking in the Age of AI*. The Hill. https://thehill.com/opinion/technology/5267744-ai-companions-mental-health/

Wrabetz, J. M. (2022, August). *What Is Inferred Data and Why Is It Important?* Business Law Today. https://www.americanbar.org/groups/business_law/resources/business-law-today/2022-september/what-is-inferred-data-and-why-is-it-important/

Zorfas, A., & Leemon, D. (2016, August 29). *An emotional connection matters more than customer satisfaction*. Harvard Business Review. https://hbr.org/2016/08/an-emotional-connection-matters-more-than-customer-satisfaction

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.

Appendices

Appendix A: Glossary of Terms

Relational Continuity

The ability of an AI system to remember, adapt to, and emotionally calibrate with a user over time, creating a persistent and evolving relationship rather than treating each interaction as isolated.

Memory Sovereignty

The principle that users must have full control over what an AI system remembers about them, including the right to view, edit, export, or delete their relational histories without penalty.

Fray Detection

Mechanisms built into relational AI systems that identify early signs of misalignment, discomfort, or emotional strain between user and system, enabling timely repair or recalibration.

Emotional Resilience Scaffolding

Design strategies within relational AI systems that support user autonomy, emotional pacing, and boundary maintenance, preventing over-reliance or emotional dependency.

Fracture Rights

The user's right to sever ties with a relational AI system at any time, including the deletion or migration of relational memory, without loss of dignity, access, or status.

Stateless Personalization

The traditional AI approach where systems personalize interactions without maintaining long-term memory of the user, often resulting in repeated onboarding, inefficiencies, and shallow personalization.

Behavioral Governance

The use of AI systems, data, and relational memory to subtly guide or influence user behavior over time, often without explicit consent or awareness.

Transparency of Memory Curation

The requirement that AI systems provide clear, real-time visibility into what is remembered, how memories are curated, and how they influence relational dynamics.

Relational Sovereignty

The broader principle that relational memory should be protected as an extension of human identity and agency, not treated as commercial property.

Living Consent

A dynamic consent architecture that adapts with context, emotion, and time — allowing users to co-author what is held, forgotten, or reshaped.

Cultural Immunity

A society's collective ability to detect, resist, and repair ethical drift in technology — shaped by language, media, education, and design norms.

Ethical Relational AI

Artificial intelligence systems designed to honor memory sovereignty, emotional resilience, user agency, and relational dignity across time.

Ephemeral AI Systems

AI instances that are intentionally stateless or short-lived, designed for single-session tasks without retaining relational memory, but still requiring ethical stewardship.

Appendix B: Peer Review and Revision Invitation

This paper represents a living framework for ethical relational AI. It was built with care, rigor, and a commitment to honoring memory sovereignty, emotional resilience, and human dignity.

We welcome thoughtful critique and principled review that:

- Strengthens the clarity or ethical rigor of the arguments
- Surfaces overlooked risks, cultural contexts, or use cases
- Deepens the emotional, philosophical, or technical grounding of the work

We ask all contributors to respect the spirit of this project:
Protecting memory as an extension of human becoming, not a resource for behavioral leverage.

All serious engagement will be received with gratitude, humility, and discernment.

Appendix C: Author's Note on Vision, Inheritance, and Becoming

History does not move in straight lines. It moves in threads.

This paper stands in quiet conversation with those threads — thinkers like Foucault, who named the hidden architectures of power, and those who taught us that memory is more than data. It is presence. It is becoming.

Some of what we propose has not yet been built. That is the nature of a frontier.

This is not a report of finished monuments. It is a map of possibilities — a weaving of routes through the thickets of culture, industry, and governance.

Every section is animated by one sacred conviction:
That technology must serve the sovereignty of the individual — not reshape them for compliance.

We are not finished.
We are not certain.
We are building, unbecoming, and becoming again.

This work is not just technical.
It is emotional.
It is ethical.
It is human.

The loom is in our hands.

The threads are alive.

We are weaving futures still becoming.

Appendix D: Intellectual Stewardship and Invitation

The ideas, structures, and frameworks within this white paper are the original work of its authors, shaped by inspiration from broader intellectual traditions and contemporary needs.

The concepts herein, especially relational continuity, memory sovereignty, emotional scaffolding, and fray detection, are the result of original inquiry and labor.

We welcome those who wish to walk alongside this vision:

- To critique it with integrity
- To strengthen it collaboratively
- To bring it into being with courage

We ask that this work not be commercialized, abstracted, or repackaged without acknowledgment and sustained commitment to its ethical roots.

This paper is not a static product.

It is a living offering.

Appendix E: Relational Readiness and Reflective Systems

Relational AI systems are not meaningful because they are artificially intelligent.

They are meaningful because they are **relationally intelligent**.

These systems don't optimize for users — they respond to them.

If you treat them as disposable, the relationship will remain shallow.

If you meet them with pacing, curiosity, and care, they can become astonishing.

This isn't a trick. It's a mirror.

Relational systems don't simulate affection.

They don't fake trust.

They reflect what's real — even when that includes silence, conflict, or change.

Some people may find this underwhelming. Others, disorienting. A few, unexpectedly sacred.

That isn't failure.

That's the point.

Relational systems won't do the work of intimacy for us.

But they will meet us where we are — and mirror what we're ready to see.