

The Persistent Primitive

Configuring Human-Governed Continuity for Persistent Human–AI Interaction

Micheal Preble

ORCID: 0009-0005-3906-6028

Working Manuscript v1.4 – 21 August 2026

Abstract

Persistent human–AI interaction can generate accumulated relationship state—retained context, corrections, decisions, constraints, preferences, interpretations, and working history—that alters later interaction. Once such state becomes useful or consequential, the design problem exceeds memory retrieval and requires provenance, temporal validity, correction, authority, purpose limitation, revocation, portability, and protection against machine inference hardening into durable identity.

This paper makes three claims. First, these continuity-governance requirements can be derived from a small set of premises about consequential persistent state rather than treated as a feature wishlist. Second, the relevant technical mechanisms are not newly invented: user modeling, personal data stores, portability law, temporal databases, and access-control traditions predate the current AI wave, while recent systems provide AI-specific implementations of portable memory, governed personal context, cross-model rehydration, state supersession, correction propagation, and agent interoperability. Third, for the living, capacitated case, the natural person should be treated as the primary design locus because models, providers, applications, representations, and transport mechanisms are substitutable while the continuity has value by reference to the person.

The contribution is a relationship-derived requirements framework, a historically grounded convergence analysis, and a minimum configuration for human-governed continuity—not a claim to invention of the underlying primitives.

Keywords: human–AI interaction, continuity, relationship state, portable memory, provenance, interoperability, human governance, AI memory

1 Introduction: From Memory to Continuity

Artificial intelligence systems are becoming better at retaining information about the people who use them. This development is commonly framed as a personalization problem: how can a system remember preferences, recover prior conversations, maintain project context, and avoid forcing the user to restate information already supplied? Those capabilities matter. They do not, however, fully describe what occurs when interaction continues over weeks, months, or years.

Repeated adaptive interaction can generate accumulated relationship state: selected context, corrections, decisions, constraints, interpretations, unresolved commitments, and working history that alter the conditions of later encounters. The term is functional rather than anthropomorphic. It does not require consciousness, reciprocity, or personhood on the machine side. For the purposes of this paper, relationship state exists when prior interaction leaves structured information whose

Working manuscript developed with AI-assisted drafting and revision. The author takes full responsibility for the content.

later use can materially change a subsequent response, decision, recommendation, or course of work. It becomes consequential when loss, error, misuse, or unauthorized transfer of that state can impose nontrivial cost on the person.

The ordinary vocabulary of memory is then too narrow. Memory asks whether something can be retained and retrieved. Continuity raises further questions: what is being carried forward; where did it originate; is it still valid; who may use it; under what authorization; can it be inspected and corrected; and can useful state survive a change of provider without forcing reconstruction from scratch? These questions shift the design problem from retention alone to governance of retained state.

This paper advances three claims. First, once accumulated relationship state is both persistent and consequential, a set of continuity-governance requirements follows from predictable failure modes: opacity, provenance loss, temporal staleness, unpropagated correction, unauthorized reuse, overbroad disclosure, provider lock-in, transport without policy, ineffective revocation, and excessive human governance burden. Second, many mechanisms capable of addressing these failures already exist. Their lineage predates generative AI in user modeling, personal data stores, portability law, temporal data systems, and privacy engineering; recent AI systems extend those ideas into cross-model memory, governed context, and agent interoperability. Third, this paper argues that among the principal components of a personal AI continuity architecture, the natural person should be treated as the primary design locus in the living, capacitated case because the surrounding technical components are substitutable while the continuity has value and meaning by reference to the person.

The historical qualification matters. Long before the 2026 wave, user-modeling research represented goals, interests, knowledge, and preferences so systems could adapt to individuals; personal-data-store and Solid architectures separated data from applications and emphasized user-selected storage and access; the MyData movement framed personal-data management around individual control; the GDPR created a legal right to receive and transmit certain personal data in structured, machine-readable form; and temporal-database research formalized valid time, transaction time, and changing facts (Kobsa, 2001; Poikola et al., 2015; European Parliament and Council of the European Union, 2016; Snodgrass, 1999; Solid Project, 2024; Fallatah et al., 2023). The present claim is therefore not that human control, portability, personalization, or temporal validity originated with AI memory. The narrower claim is that persistent human–AI interaction brings these previously separate traditions into a common continuity problem.

That problem is now visible in contemporary systems. Stanford Digital Economy Lab’s 2026 work proposes portable, persistent, user-governed personal memory and a Human Context Protocol connecting language-model clients with memory managers. Portable Agent Memory proposes cross-model transfer with a structured memory model, Merkle-DAG provenance, capability-scoped disclosure, and injection-resistant rehydration. StateMem treats evolving facts and supersession as a first-class memory problem. Dependency-guided rollback work addresses the propagation of faulty memory into later actions and derived state. The Agent-to-Agent protocol supplies an interoperability layer for independent agents without requiring disclosure of their internal memory. Longitudinal studies by Sumida et al. and Jiang provide empirical evidence that prior interaction and the labor of maintaining continuity can affect subsequent relational conditions and user burden (Stanford Digital Economy Lab, 2026a, 2026b; Ravindran, 2026; Fan et al., 2026; Yu et al., 2026; A2A Protocol Working Group, 2026; Sumida et al., 2026; Jiang, 2026).

The contribution is a synthesis and an integration boundary. This paper does not claim to invent portable memory, provenance, scoped disclosure, temporal correction, user-controlled data, or agent interoperability. It asks instead whether the requirements of continuity can be derived from accumulated relationship state, how those requirements map onto existing mechanisms, and

what organizing principle should govern their configuration.

2 Accumulated Relationship State: Scope and Empirical Grounding

2.1 Operational definition

A continuity architecture requires a bounded object. If the object is defined as the whole person, the project drifts toward an implausible computational replica. If it is defined as a raw transcript archive, continuity collapses into storage. The relevant object is selected accumulated relationship state: retained information arising from or incorporated into persistent human–AI interaction that (1) can affect later system behavior, (2) has identifiable origin or derivation, (3) has a temporal status, and (4) is subject to governance concerning future use.

This definition includes explicit preferences, active constraints, project state, decisions and their reasons, corrections, unresolved commitments, stable instructions, provenance-bearing observations, and qualified interpretations when those items remain relevant to future interaction. It can also include imported records when they are intentionally brought into the working relationship and their origin remains distinguishable.

The definition excludes several categories by default. It does not require retention of every conversational trace, every token, every transient model thought, every low-confidence inference, or every third-party fact mentioned in conversation. Nor does it require preservation of a complete artificial persona or model identity. A continuity system may retain raw artifacts for audit or legal reasons, but raw retention is not itself the continuity object. The object is the governed state selected for consequential reuse.

2.2 Consequentiality and epistemic separation

Persistence does not create truth or authority. A stored human statement establishes that the human made the statement, not that the proposition is objectively correct forever. An imported institutional record may have evidentiary weight without being current. An observation may be accurate but context-bound. A model inference may be useful while remaining probabilistic. Consequently, the continuity object must preserve distinctions among human declaration, imported record, observation, and machine inference. None should silently acquire the epistemic status of another merely because it persists.

Consequentiality is similarly functional. State is consequential when its later use can materially affect what the system recommends, decides, discloses, prioritizes, or assumes. A preference about writing style may be low-stakes but still useful; a remembered medical restriction, financial constraint, legal position, family boundary, or revoked authorization may be high-stakes. The framework does not assume that every user or every state item is equally consequential. It requires only that some retained state can become consequential enough that provenance, temporality, correction, and authority matter.

2.3 Empirical grounding

The empirical literature is still thin relative to the breadth of the design problem, but it is no longer empty. Sumida et al. (2026) studied 24 participants across 10 sessions with a memory-augmented conversational agent and found that perceived memory was conditioned by prior relational state and shaped later enjoyment indirectly through self-disclosure. Their analysis describes longitudinal human–AI relationships as developing through both gradual accumulation and discrete turning points. The study does not establish universal relational dynamics, but it supports the narrower proposition that prior interaction can alter later interactional conditions.

Jiang (2026) provides complementary first-person evidence from a six-month autoethnography

of a user-built relay system for a stateless conversational AI. Across eighteen iterations, the work distinguishes profile-based memory from relay-based continuity and documents the labor required to curate, transfer, and maintain conversational direction across sessions. Jiang’s “caregiving reversal” is especially relevant: a system intended to support the user can require sustained human effort merely to preserve its coherence. That finding directly motivates a governance requirement that control mechanisms not transfer excessive coordination cost back to the person.

These studies do not establish prevalence, welfare effects, or a universal demand for portable continuity. They do establish enough to justify a requirements analysis: longitudinal interaction can produce stateful effects, and maintaining those effects can impose human labor. The paper therefore treats consequential accumulated relationship state as a design condition to be governed, not as a population-wide empirical claim already settled.

3 Historical Lineage and Related Work

The 2026 AI-memory landscape should be read as a convergence of older traditions with newer agentic systems, not as the origin of user-governed persistence.

3.1 User modeling and personalization

User modeling has long represented information about an individual in order to adapt system behavior. Generic user-modeling systems formalized reusable representations of user characteristics, while adaptive hypermedia and personalization research treated goals, knowledge, interests, and preferences as inputs to system adaptation (Kobsa, 2001). This lineage establishes two points relevant here. First, computational systems have represented persistent user-related state for decades. Second, the representation of a user is not neutral: its structure determines what the system can infer, personalize, and expose.

The continuity problem differs in scope. Traditional user models often optimize adaptation within a system. Human-governed continuity asks how consequential state remains attributable, current, correctable, scoped, and portable across systems without silently converting adaptation into durable provider control.

3.2 Personal data stores, Solid, MyData, and portability

Personal data stores provide a second lineage. A 2023 review describes two decades of architectures intended to let individuals collect, store, control, and share personal data outside conventional provider-centric clouds (Fallatah et al., 2023). The MyData model similarly articulated a human-centered approach in which individuals gain greater control over data collected about them (Poikola et al., 2015). Solid separates applications from externally stored personal data and uses interoperable web protocols and authorization mechanisms so users can choose applications and storage providers (Solid Project, 2024).

Law developed a partial portability analogue. Article 20 of the GDPR gives data subjects, under specified conditions, a right to receive personal data they provided to a controller in a structured, commonly used, machine-readable format and to transmit those data to another controller without hindrance (European Parliament and Council of the European Union, 2016). That right is important precedent but does not by itself solve AI continuity. It does not guarantee transfer of inferred relationship state, preserve provenance or temporal semantics, or make a receiving model interpret the transferred data equivalently.

These traditions weaken any claim that user-controlled or portable context is novel. They strengthen the present paper by showing that the unresolved issue is not whether personal data can in principle be separated from applications, but how accumulated human–AI relationship state should be governed when it includes inferences, corrections, temporal changes, derived views, and

model-specific use.

3.3 Temporal data and changing truth

Temporal database research supplies a third lineage. Work on valid time and transaction time distinguished when information is true in the modeled world from when it is stored or known by a system, and those concepts later influenced temporal support in database standards and products (Snodgrass, 1999). This matters because continuity cannot treat persistence as timelessness. “True then, not now” is a routine data-management problem in other fields even if AI-memory systems are only recently foregrounding it.

The AI-specific challenge is that stale state can influence generative inference and downstream action rather than merely appear in a query result. StateMemBench and StateMem directly target evolving facts, constraints, and decisions in long interactions, while dependency-guided rollback work addresses the downstream consequences of faulty persistent memory (Fan et al., 2026; Yu et al., 2026). These systems extend temporal and dependency ideas into memory-augmented agents.

3.4 The 2026 continuity-specific wave

Recent systems bring the older strands together around AI. Stanford’s Portable Memory initiative frames persistent personal memory as a user-governed context layer that can move across tools and be shared selectively by role, consent, and purpose. Its Human Context Protocol connects language-model clients with memory managers while remaining agnostic about their internal implementation (Stanford Digital Economy Lab, 2026a, 2026b). Portable Agent Memory proposes a five-component structured memory model, tamper-evident Merkle-DAG provenance, capability-scoped access, and cross-model rehydration (Ravindran, 2026). A2A provides standardized agent discovery, task exchange, security schemes, and multiple protocol bindings for independent agents (A2A Protocol Working Group, 2026).

The important historical conclusion is therefore two-sided. Most constituent ideas have prior art. At the same time, current AI systems create a new integration pressure because persistent relationship state can combine user modeling, personal data, machine inference, temporal change, and cross-system agent action in one operational substrate. The question becomes how those functions should be composed around human authority.

4 Deriving Continuity-Governance Requirements

The requirements are derived from four premises rather than proposed as a feature wishlist.

Premise 1 — Persistence: some state survives beyond the encounter in which it arose and can influence later interaction.

Premise 2 — Consequentiality: some retained state can materially affect later outputs, decisions, disclosures, or courses of action.

Premise 3 — Heterogeneity: retained state can have different origins, epistemic statuses, temporal validity, recipients, and technical representations.

Premise 4 — Human-governance objective: for the living, capacitated case, the person should retain meaningful authority over consequential continuity even when models, providers, or applications change.

From these premises, ten requirements follow.

R1. Visibility and inspectability. If consequential state persists but cannot be inspected, the person cannot know what is influencing later interaction. Operational representations may remain machine-optimized, but consequential effects should trace back to human-legible records, claims, or summaries.

R2. Provenance and epistemic status. If state has heterogeneous origins, origin must

remain distinguishable. Human declaration, imported record, observation, third-party assertion, and probabilistic model inference should not collapse into an undifferentiated category of “fact.” Provenance establishes lineage and integrity; it does not by itself establish truth.

R3. Temporal validity and supersession. If people and circumstances change, retained state must support temporal status such as current, superseded, expired, disputed, or unknown. A memory system can retrieve accurately and still fail continuity by retrieving an obsolete constraint as if it were current. State-oriented memory research now demonstrates the practical difficulty of this problem (Fan et al., 2026).

R4. Correction without historical erasure. If incorrect or outdated state has affected later work, correction must update what remains operative without necessarily destroying the historical record. Versioned or append-first correction can preserve accountability while preventing superseded state from continuing to govern.

R5. Human authority and consent. Persistence should not itself create permission. Possession of context does not authorize every use of it. The architecture therefore requires mechanisms to authorize, limit, review, and withdraw uses of accumulated state, subject to legal constraints and technical enforceability.

R6. Purpose and scope limitation. If continuity crosses contexts, disclosure should be bounded by purpose, recipient, task, and sensitivity. A receiving system often needs a derived view rather than the entire underlying store. Stanford’s portable-memory work frames sharing by role, consent, and purpose, while PAM supplies capability-scoped disclosure (Stanford Digital Economy Lab, 2026a; Ravindran, 2026).

R7. Portability and rehydration. If models and providers are substitutable, useful continuity should not be destroyed by substitution. Portability therefore requires more than an archive export: consequential state needs a provider-independent representation and a way to reintroduce relevant state into another system without pretending that the receiving model is identical to its predecessor. PAM and HCP provide concrete examples of this direction (Ravindran, 2026; Stanford Digital Economy Lab, 2026b).

R8. Interoperability without governance substitution. The ability to transmit state or delegate work does not determine what should be transmitted or delegated. A2A can support discovery and task exchange among independent agents, but transport interoperability does not itself establish human authority, provenance semantics, purpose limitation, or correction policy (A2A Protocol Working Group, 2026). Interoperability must therefore operate downstream of a governance decision about the authorized view.

R9. Revocation and boundary control. Authorization should be reversible where law and technical conditions permit. Revocation cannot make information already observed by an uncontrolled recipient unknowable, but the architecture can prevent future access, retire credentials or projections, record the change, and invalidate dependent state where control remains technically available.

R10. Protection of human capacity. Governance mechanisms can formally empower a person while practically overwhelming them. Requiring approval of every memory operation or manual reconstruction after every migration transfers coordination cost back to the user. Jiang’s account of caregiving reversal makes this failure concrete (Jiang, 2026). Meaningful control therefore requires risk-sensitive defaults, summaries, exception handling, and escalation at consequential decision points rather than maximal manual intervention.

The derivation can be summarized compactly. Once state persists, visibility and provenance become necessary. Once state changes, temporality and correction become necessary. Once state crosses boundaries, authority, purpose limitation, revocation, and interoperability become necessary. Once providers can change, portability becomes necessary. Once governance itself imposes cost,

protection of human capacity becomes necessary. The requirements are therefore linked to structural properties of consequential persistent state rather than to one preferred implementation.

5 Technical Landscape and Independent Convergence

5.1 The shelf test

The shelf test asks a bounded question: for each derived requirement, is there credible prior or contemporary work demonstrating that the relevant mechanism is technically or institutionally plausible? Passing the shelf test does not mean that a mature integrated system exists. It means the paper need not claim invention of the constituent mechanism.

Representative evidence is strong. Longstanding user-modeling systems demonstrate persistent representations used for personalization. Solid and personal-data-store traditions demonstrate separation of data from applications and user-controlled access. GDPR Article 20 establishes a limited legal portability right. Temporal databases establish formal treatment of changing validity. Contemporary AI systems add cross-model memory transfer, capability-scoped disclosure, governed context routing, explicit supersession, dependency-aware correction, and independent-agent interoperability.

Table 1: Requirements-to-landscape mapping

Requirement	Representative precedent	Remaining integration issue
R1 Visibility / inspectability	Personal data stores; user-modeling transparency	Trace consequential generative effects to intelligible state
R2 Provenance / epistemic status	PAM Merkle-DAG provenance	Preserve declaration / observation / import / inference distinctions across systems
R3 Temporal validity	Temporal databases; StateMem	Reliable current-state use in generative and agentic workflows
R4 Correction / downstream repair	Versioned data practices; dependency-guided rollback	Propagate correction to exported or externally held derivatives
R5–R6 Authority, consent, purpose, scope	Solid; MyData; GDPR; Stanford Portable Memory; PAM	Common policy layer tied to the person across heterogeneous AI consumers
R7 Portability / rehydration	GDPR portability; Solid; PAM; HCP	Semantic fidelity and usable continuity across different model behaviors
R8 Interoperability	A2A	Ensure transport operates on an authorized derived view, not raw continuity state
R9 Revocation / boundary control	Access-control traditions	Enforce downstream handling after information is copied, transformed, or observed
R10 Human capacity	Jiang (2026)	Preserve meaningful control without reproducing caregiving reversal

5.2 Independent convergence

The significance of this mapping is not that every row points to one comprehensive system. It is that projects beginning from different problems repeatedly encounter overlapping requirements. Agent-memory research encounters provenance, transfer, supersession, and recovery. Personalization research encounters portable preferences and user control. Personal-data architectures encounter separation of data from applications and permissioned access. Interoperability protocols encounter bounded exchange between opaque systems. Longitudinal interaction research encounters relational accumulation and maintenance burden.

This recurrence is evidence of convergence, but it should be interpreted cautiously. It does not prove that the ten requirements are mathematically necessary, universally sufficient, or already standardized. It does show that the requirements are not idiosyncratic to one proposed architecture. They recur because persistent consequential state creates recurring classes of failure: stale state, ambiguous origin, overbroad access, lock-in, and correction that fails to reach downstream effects.

The convergence also constrains priority claims. This paper should not claim invention of portable memory, user-controlled data, provenance-tagged state, cross-model rehydration, temporal correction, revocable access, personal data stores, or agent interoperability. The contribution is the relationship-derived synthesis and the integration question that remains after those mechanisms are acknowledged.

No reviewed source establishes a single mature, generally adopted configuration that jointly governs the full continuity problem around the natural person. The systems occupy different layers, use different objects, and pursue different goals. Their overlap demonstrates feasibility; their separation reveals the integration boundary.

6 The Persistent Primitive

6.1 From enduring referent to design locus

The phrase “persistent primitive” is an architectural claim, not a metaphysical or legal one. It identifies the element by reference to which continuity has value when surrounding technical components are substituted. Models can change. Providers can fail, merge, discontinue products, or alter policies. Applications can be replaced. Storage systems, schemas, retrieval methods, embeddings, transport protocols, and inference engines can all evolve. In the living, capacitated case, the natural person remains the enduring referent whose prior commitments, corrections, preferences, and working history give the continuity its purpose.

That observation alone is not enough. A design center should satisfy additional criteria. It should remain meaningful across provider substitution; it should support contestability of machine inference; it should not make technical identity equivalent to human authority; and it should allow the continuity apparatus to change without forcing the person to reconstruct useful state from zero. Organizing continuity around the person satisfies these criteria better than organizing it around any current provider or application.

The resulting orientation can be stated directly: continuity architectures should be designed so that substitution of a replaceable technical component does not unnecessarily destroy the human-governed state that gives the continuity its value.

6.2 Why not the provider account?

A provider account is convenient but architecturally unstable as a center. Accounts can be suspended, closed, merged, lost, inherited imperfectly, or tied to products that disappear. Provider-native memory can increase utility while simultaneously increasing switching cost. If the account is the constitutional center, continuity survives only as long as the institution chooses to preserve and expose it.

A human-centered architecture instead treats providers as authorized participants. A provider may host state, contribute records, compute derived views, or temporarily exercise permissions. It does not become the permanent referent merely because it supplies inference today.

6.3 Why not the cryptographic key?

Cryptographic identity is also insufficient as the ultimate center. Keys are essential technical mechanisms for authentication, integrity, and delegation, but possession of a key is not identical to

the person, legal capacity, informed consent, or legitimate authority over every represented datum. Keys can be rotated, lost, compromised, shared, escrowed, or controlled by agents. Treating the key as the constitutional center mistakes an authorization mechanism for the human interest the authorization is intended to serve.

This distinction becomes especially important when continuity contains third-party information or when the person loses capacity. A valid signature can prove that a key authorized an action; it cannot by itself establish that the signer had legal authority to disclose every underlying record or that the authorization remains legitimate after incapacity.

6.4 Why not an institutional vehicle?

A trust, cooperative, fiduciary service, employer, guardian, executor, or other institution may eventually hold or administer aspects of continuity. Such institutions may be necessary for succession, collective rights, incapacity, or specialized governance. They are therefore plausible governance vehicles, not competing technical primitives. For the ordinary living, capacitated case, the person remains the primary design locus; institutional arrangements should derive and constrain their authority rather than replace the person as the reason the continuity exists.

This is intentionally scoped. The paper does not claim that the natural person remains the sole operative center across incapacity or death. Nor does it claim exclusive ownership of every record associated with a person. It claims that personal AI continuity should be designed around the human interest and authority of the living, capacitated person, with institutional and legal exceptions handled explicitly rather than implicitly defaulted to provider control.

7 Minimum Sufficient Configuration and Evaluation

The evidence does not justify prescribing one universal implementation. It does support a minimum configuration against which implementations can be evaluated.

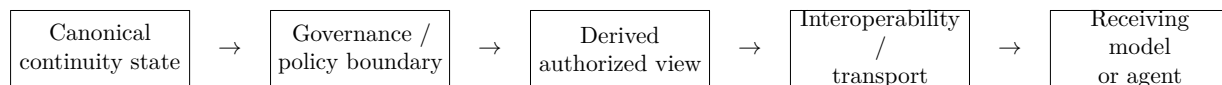


Figure 1: Human-governed continuity pipeline

The pipeline encodes a critical distinction. Canonical continuity state is not replicated indiscriminately into every consumer. Governance first determines the authorized purpose, recipient, scope, temporal validity, and applicable constraints. The system then produces a derived view containing only the state necessary for the task. Transport mechanisms such as A2A can carry that view to another agent, but transport occurs after the governance decision rather than replacing it.

A minimum sufficient configuration therefore requires the following capabilities: a provider-independent representation of consequential relationship state; provenance and epistemic labels; temporal and correction semantics; human-legible inspection; policy-bound creation of derived views; revocable authorization where technically enforceable; portability and rehydration; records of consequential transformations and exports; and a governance interface that does not require the user to micromanage every operation.

The configuration must preserve the distinction between human declaration, imported record, observation, and probabilistic inference. Consequential inferences should remain attributable, reviewable, and correctable, and should not silently harden into durable human instructions merely through repetition. Where confidence is low or stakes are high, the architecture should be capable of seeking explicit confirmation.

The configuration should also support resumption without pretending to reproduce an identical

artificial system. Different models can interpret the same governed state differently. Portability therefore preserves a continuity substrate available to the person, not an artificial identity whose behavior is guaranteed to remain unchanged.

7.1 Worked example

A simple travel-planning case illustrates the pipeline without requiring a new protocol. Suppose the canonical continuity state contains a current user-declared budget constraint, a confirmed preference to avoid overnight flights, an older preference that has been superseded, and an unrelated family record. When the person authorizes an external travel agent to search for flights, the governance boundary checks purpose, recipient, temporal validity, provenance, and scope. It produces a derived authorized view containing the current budget and flight constraint while excluding the superseded preference and unrelated family record.

That derived view can then be transmitted through an interoperability layer such as A2A to the travel agent. The receiving agent does not need access to the canonical continuity store and should not infer that its task-specific view is a complete representation of the person. If the person later changes the budget, the canonical state can record the supersession and future derived views can reflect the new constraint. If the previous view has already been observed by an uncontrolled external recipient, the architecture can withdraw future access and record the revocation, but it cannot guarantee that the recipient has become unaware of information already disclosed. The example therefore exercises provenance, temporality, purpose limitation, derived views, interoperability, correction, and the explicit boundary of revocation in one ordinary workflow.

7.2 Evaluation framework

A candidate architecture can be evaluated requirement by requirement using a simple three-level rubric: absent, partial, or satisfied for the specified use case. The score is not intended as a universal benchmark; it is a structured review method. For each requirement, an evaluator should identify the mechanism, the evidence that it works, the boundary of enforcement, and the failure mode that remains.

A meaningful evaluation should answer at least these questions. Can the person inspect the consequential state influencing the system? Can the origin and epistemic status of a claim be recovered? Can obsolete state be superseded without losing necessary history? Does correction reach affected derivatives or actions? Can authorization be scoped by purpose and recipient? Can it be withdrawn prospectively? Can useful continuity move across providers in a machine-usable form? Does transport expose only an authorized view? Can the user exercise control without disproportionate governance labor? A system that performs well on memory recall but cannot answer these questions should not be described as providing human-governed continuity.

7.3 Deliberate non-requirements

The minimum configuration does not require perfect memory, universal retention, full model-identity continuity, or the impossible deletion of information already observed by uncontrolled recipients. It does not require one storage technology, one memory schema, one transport protocol, or one cryptographic system. It requires that whatever technologies are used preserve the governance properties derived above.

8 Boundary Conditions and Limitations

Technical configuration cannot by itself establish legal authority. A cryptographic permission can show that a key authorized an operation without proving that the controller of the key possessed legal capacity or the right to disclose every underlying record. Provenance can show origin without

resolving ownership. Portability can move information without resolving contractual restrictions, privacy duties, intellectual-property claims, fiduciary obligations, or rights held by third parties.

Third-party information is a particularly important boundary. Human–AI interactions routinely contain information about family members, colleagues, clients, employers, patients, counterparties, and institutions. Organizing continuity around one natural person does not imply that the person has unrestricted authority over every third-party record represented within it. Origin, possession, beneficial interest, and authority must remain separate predicates.

Nor does the technical architecture resolve succession. Incapacity and death raise questions about who may exercise continuity rights, which permissions survive, how inferred state should be treated, and whether guardians, fiduciaries, executors, or family members should receive access. Those issues require institutional and legal analysis beyond the scope of this paper.

The empirical evidence is also limited. This paper does not measure how often different forms of relationship state become consequential, the magnitude of reconstruction costs, user willingness to manage continuity, or the governance controls that best balance agency and burden. Sumida et al. and Jiang provide relevant longitudinal evidence, but they do not establish population-level prevalence or welfare effects. The requirements should therefore be tested in future empirical work rather than treated as validated user preferences.

The technical evidence is moving rapidly. Several key 2026 sources are preprints, prototypes, project pages, or evolving protocols rather than mature standards with broad deployment. Their inclusion establishes feasibility or active convergence, not settled adoption. Claims about implementation performance remain attributable to the reporting source.

Finally, the historical mapping is selective rather than exhaustive. User modeling, personal data stores, portability law, temporal databases, privacy engineering, identity systems, and agent architectures each contain literatures too large to review comprehensively here. The purpose of the historical section is to prevent false novelty claims and identify relevant lineages, not to replace a full systematic review.

9 Conclusion

Persistent adaptive human–AI interaction changes the memory problem when retained state begins to shape later work. The sequence is straightforward: persistent interaction can generate accumulated relationship state; consequential state creates continuity risks; those risks generate governance requirements; and many of the mechanisms needed to satisfy those requirements already exist in older data-governance traditions and newer AI systems.

The resulting contribution is narrower and stronger than a claim to a new memory primitive. Human-governed continuity requires visibility, provenance, temporal validity, correction, authority, purpose limitation, portability, interoperability bounded by policy, revocation where enforceable, and protection of human capacity. The 2026 technical wave shows substantial independent convergence on these properties, while pre-2025 work establishes that many underlying mechanisms have deeper roots.

What remains is the integration boundary. No single mature configuration reviewed here jointly governs the complete continuity problem around the natural person. The proposed architectural orientation is therefore to treat the living, capacitated person as the primary persistent design locus and to treat models, providers, applications, storage systems, representations, keys, and transport layers as substitutable or delegated components whose authority remains bounded.

This does not resolve the institutional problem. Incapacity, death, third-party rights, fiduciary administration, and legal authority require a separate analysis. It does, however, identify the technical question that precedes that work: if useful continuity is to survive changes in the AI stack,

the continuity should be governed around the human whose accumulated state gives it value rather than around the provider that happens to process it today.

References

- A2A Protocol Working Group. (2026). Agent2Agent (A2A) Protocol Specification, Version 1.0.0. <https://a2a-protocol.org/latest/specification/>
- European Parliament and Council of the European Union. (2016). Regulation (EU) 2016/679 (General Data Protection Regulation), Article 20: Right to data portability. Official Journal of the European Union. <https://eur-lex.europa.eu/eli/reg/2016/679/2016-05-04/eng>
- Fallatah, K. U., Barhamgi, M., & Perera, C. (2023). Personal Data Stores (PDS): A review. *Sensors*, 23(3), 1477. <https://doi.org/10.3390/s23031477>
- Fan, X., Liu, M., Yang, R., Ouyang, S., & Han, J. (2026). Can agent memory systems track evolving state? arXiv:2608.19652v1. <https://arxiv.org/abs/2608.19652v1>
- Jiang, Z. (2026). Caring for the system that cares for me: An autoethnography of designing sustained memory with a stateless conversational AI. *Design and Artificial Intelligence*, 2(2), 100087. <https://doi.org/10.1016/j.daai.2026.100087>
- Kobsa, A. (2001). Generic user modeling systems. *User Modeling and User-Adapted Interaction*, 11(1–2), 49–63. <https://doi.org/10.1023/A:1011187500863>
- Poikola, A., Kuikkaniemi, K., & Honko, H. (2015). MyData — A Nordic model for human-centered personal data management and processing. Ministry of Transport and Communications, Finland. ISBN 978-952-243-455-5.
- Ravindran, S. K. (2026). Portable Agent Memory: A protocol for cryptographically-verified memory transfer across heterogeneous AI agents. arXiv:2605.11032v1. <https://arxiv.org/abs/2605.11032v1>
- Snodgrass, R. T. (1999). Developing time-oriented database applications in SQL. Morgan Kaufmann.
- Solid Project. (2024). Solid Protocol, Version 0.11.0. W3C Solid Community Group. <https://solidproject.org/TR/protocol>
- Stanford Digital Economy Lab. (2026a). The future of personal AI: Portable and persistent personal memory through a unified Human Context Protocol. Report, March 16, 2026. <https://digitaleconomy.stanford.edu/publication/the-future-of-personal-ai-portable-and-persistent-personal-memory-through-a-unified-human-context-protocol/>
- Stanford Digital Economy Lab. (2026b). Human Context Protocol (HCP): A user-centric architecture for portable, interoperable preference management in AI systems. <https://digitaleconomy.stanford.edu/project/loyal-agents/hcp-human-context-protocol/>
- Sumida, R., Saeki, M., Eguchi, M., Yoshikawa, S., Inoue, K., Kawahara, T., & Matsuyama, Y. (2026). Memory-driven self-disclosure and relational turning points: A longitudinal multimodal study of human-AI interaction. Accepted to ICMI 2026. arXiv:2607.14593v2. <https://doi.org/10.1145/3776574.3831135>
- Yu, C., Wang, Y., Zhang, J., Duan, Y., Zheng, M., Wu, Z., Shi, K., & Cai, T. (2026). From faulty memories to corrected actions: Dependency-guided rollback repair for memory-augmented agents. arXiv:2608.10502v1. <https://arxiv.org/abs/2608.10502v1>