From Preservation to Reachability:
Semantic Continuity in an Age of Drift

Flyxion

Independent Researcher

`github.com/standardgalactic`

April 17, 2026

Abstract

Contemporary information systems are routinely described as preserving knowledge: archives store documents, encyclopedias consolidate pages, citation systems maintain references, and platforms retain audiences. This paper argues that preservation is the wrong invariant. What matters for the continuity of meaning over time is not whether a semantic state exists somewhere in a system, but whether that state remains *reachable*—navigable from present states through available transformations. We formalize this distinction, exhibit five canonical failure modes in which preservation succeeds while reachability collapses, and extend the analysis to linguistic systems operating at longer timescales. The paper concludes by sketching a constructive requirement: systems that aim to sustain semantic continuity must preserve not only nodes but the paths between them, a condition that connects naturally to constraint-based accounts of meaning under transformation.

---

1 Introduction

A library that burns is understood as a catastrophe. A library whose catalog is deleted, whose call numbers are reassigned, and whose shelves are rearranged without documentation is understood as— at worst—a management failure. Yet in both cases the result is the same: a reader arriving with a reference cannot find the book. The second library has *preserved* its collection while destroying the conditions under which the collection is useful.

This contrast captures the central problem of the present paper. Modern digital and institutional systems are organized around preservation: the persistence of files, pages, references, and audiences. But preservation, in the sense of continued existence, is neither necessary nor sufficient for what we actually require from memory systems, which is *reachability*: the ability of agents in future states to arrive at past semantic content through available transformations.

The distinction matters because preservation and reachability can fail independently. A state can be preserved but rendered unreachable (the archive that no search engine indexes). A path can exist but lead nowhere (the link that resolves to a renamed page with overwritten content). And a system can maintain all its original components while gradually eliminating the connections among them.

The argument proceeds in four stages. Section 2 introduces the formal distinction between preservation and reachability in minimal terms. Section 3 presents a taxonomy of five failure modes—each a distinct way the two can come apart—drawn from concrete cases in digital infrastructure and knowledge management. Section 4 extends the analysis to natural language systems, where the same structure operates at the scale of generations rather than server cycles. Section 5 sketches the constructive implication: what a system must do if it aims to preserve reachability rather than merely storage.

A single formulation crystallizes the distinction before the formalism: a system can *remember perfectly* while being *unable to recall*. Memory, in this sense, is the existence of a state in the system. Recall is the existence of a path from here to there. The two are independent, and conflating them is the source of most design failures in knowledge infrastructure.

2 Preservation and Reachability

Let a *semantic system* consist of a set of states $\mathcal{S}$ and a set of transformations $\mathcal{T}$ that agents can apply to move between states. These transformations include any operation that constitutes legitimate access or navigation in the system: following a hyperlink, querying a search index, looking up a citation, using a translation, recalling from a prior interaction. The pair $(\mathcal{S}, \mathcal{T})$ induces a directed graph $\mathcal{G}$ in which states are nodes and transformations are edges.

Preservation is the property that a state $s \in \mathcal{S}$ continues to exist over time—that is, that the node remains in the graph.

Reachability is the property that a state s remains in the forward closure of the present state s_0 under available transformations—that is, that there exists a directed path from s_0 to s in $\mathcal{G}$.

The central claim of this paper is:

> *A system can preserve elements of $\mathcal{S}$ while destroying the path structure of $\mathcal{G}$, such that preserved states become unreachable from the present.*

This claim is not a logical paradox but a structural observation about how systems are actually built. Nodes and edges are maintained by different mechanisms—storage systems versus indexing and linking systems—and they degrade through different processes. Preservation is addressed by backup policies, archival standards, and data redundancy. Reachability is addressed by navigation infrastructure: search, citation, cross-referencing, naming consistency. Because these are separate engineering concerns, they can fail separately.

Floridi's account of semantic information treats information as meaningful and truthful data [2], but the question of *access* to information is largely treated as prior to or outside the information-theoretic frame. Shannon's channel model [1] similarly brackets semantic content in favor of transmission fidelity. What neither framework captures directly is the temporal and navigational dimension: whether a receiver in a future state can reconstruct the path to a past message through the transformations available at that future time. Reachability, as we use the term here, names this gap.

3 A Taxonomy of Discontinuities

The following cases are not chosen for comprehensiveness but for structural distinctness. The first five each represent a different way in which the reachability graph

$\mathcal{G}$ can be damaged independently of the state set $\mathcal{S}$. The sixth inverts the pattern, showing that reachability can be sustained without formal storage infrastructure at all. Together the six constitute what we will call the *minimal generating set* of cases for understanding semantic continuity.

3.1 Namespace Laundering: Identity Drift with Audience Retention

On large social platforms, named communities—groups, pages, channels— accumulate audiences over time. The audience constitutes the social infrastructure through which content is distributed and meaning is contextually stabilized. When a community is renamed or its stated purpose is changed, the audience is typically retained while the semantic identity of the community is overwritten.

This operation, which we call *namespace laundering*, severs the mapping between a name and its prior referent while conserving the distribution network. Audience members who joined under one semantic contract now participate in one that has been silently revised. New members encounter the community only in its present form, with no record that a prior form existed.

The failure here is not loss of a node but corruption of the edge label. The path from a present state to the community's prior meaning passes through a name that now points elsewhere. The prior meaning is not preserved; it is overwritten at the level of the identifier.

3.2 Encyclopedic Consolidation: Path Compression with Loss

Reference systems such as Wikipedia stabilize meaning through consolidation: articles are merged, disambiguated, or renamed as editors converge on preferred framings. This process improves navigability in one direction—from the present toward canonical content—while eliminating intermediate nodes that represented prior framings, minority perspectives, or earlier consensus states.

When an article is merged into another, the prior article's existence is recorded only in edit histories, which are not indexed by external search and require specific navigational knowledge to access. A user arriving from a search query encounters the current consensus without any indication that a prior structure existed. The prior structure is preserved in the sense that it can be reconstructed by someone who already knows to look, but it is not reachable from the present by available general-purpose transformations.

This is path compression with loss: the graph is simplified, but the simplification is irreversible from the outside.

3.3 Access Restriction: Gated Edges

Academic publishing systems preserve a large proportion of the scholarly record. Digital Object Identifiers [3] provide persistent references that survive server changes and link decay. Articles exist, are indexed, and can be cited. But without institutional subscription credentials, the edge from a reference to its referent is blocked at the point of traversal.

The node is present in $\mathcal{G}$. The edge exists. But the edge is *gated*: traversal is restricted to agents who satisfy an external condition unrelated to the semantic content of the document. From the perspective of an agent without credentials, the state is preserved but unreachable. The citation points to an object that cannot be accessed through available transformations.

3.4 Link Rot: Edge Deletion

Web infrastructure links are edges, and edges degrade. Studies of link persistence in scholarly and journalistic contexts consistently find substantial decay rates over periods of years [4]. Domains lapse, servers are decommissioned, and URL structures are reorganized without redirection. The result is a reference that resolves to nothing: a citation without an object, a pointer with no target.

This is the most straightforward case—literal edge deletion. The node may have existed; it may no longer. What is certain is that no path from the present state leads to it. The failure is not at the level of preservation policy but at the level of infrastructure maintenance: the transformations that once existed no longer do.

3.5 Archival Isolation: Node Preservation with Path Loss

Archival projects represent the most serious attempt to address preservation—the deliberate maintenance of digital states against decay. But preservation of a node is not sufficient if no paths lead to it.

The September 11 Digital Archive, maintained by the Roy Rosenzweig Center for History and New Media and held by the Library of Congress, preserves a substantial record of firsthand accounts and digital artifacts from 2001. It is not indexed by major search engines in ways that surface it for relevant queries. Users who do not already know of its existence typically discover it only through secondary references—citations in academic work, mentions in archival literature, or the Wayback Machine's own indexing.

The archive is preserved. It is not reachable from the present by the transformations available to a general-purpose user. The path exists—in principle—but requires specific navigational knowledge that is itself subject to decay.

This case reveals a deeper point: reachability is not a binary property of a system but a function of the navigational resources available to a particular class of agent at a particular time. More precisely, reachability is a property not of the graph $\mathcal{G}$ alone, but of the pair $(\mathcal{G}, A)$, where A denotes the capabilities of a given agent—the transformations that agent can actually perform. Two agents inhabiting the same system may therefore inhabit different effective graphs: the credentialed researcher and the uncredentialed reader face different subgraphs of $\mathcal{G}$, as do the trained archivist and the general-purpose search user, or the fluent speaker and the monolingual outsider. An archive reachable by specialists is not reachable in the sense that matters for broad semantic continuity.

3.6 Dense Local Systems: Reachability Without Archival Memory

The five cases above are all failure modes. It is instructive to examine a contrasting case—a system in which reachability is effectively guaranteed not by archival infrastructure but by path density.

Consider a small, stable community with high identity continuity: a neighborhood, a craft guild, a religious congregation, a family that has occupied the same region for generations. Such communities typically have weak formal archival infrastructure. They do not maintain indexed records of past decisions or digitized repositories of prior knowledge. Yet semantic continuity is often extremely robust.

The reason is structural: the reachability graph $\mathcal{G}$ is locally dense. Every present state is connected to past states through multiple redundant paths—overlapping memory among members, repeated narrative, ritual re-enactment, shared material culture, and the direct presence of individuals who participated in prior events. If one path degrades, others remain. The system does not need formal preservation because it continuously traverses its own history.

This case makes the theoretical point most cleanly: storage is not the invariant. A system with minimal storage and maximal path density may sustain semantic continuity more robustly than a system with comprehensive archives and degraded navigation infrastructure. The dense local case is not a nostalgic ideal; it is a structural limit that clarifies what the failure modes are actually failing to reproduce. What large-scale information systems lose when they scale is not primarily storage capacity—it is exactly this local redundancy of paths.

4 Language as a Long-Timescale Instance

The five failure modes described above operate at the timescale of years and decades. The same structure is visible in natural language systems, but the timescale is generations, and the mechanisms are social and political rather than technical.

Language is not simply a medium of communication; it is a navigation system for a vast inherited semantic field. A speaker of a language inherits access to a particular subgraph of human meaning—the texts, oral traditions, conceptual structures, and interpersonal knowledge encoded in that language's history. A language shift—whether voluntary or imposed—is not merely a change in expressive medium. It is a change in the available transformations: the paths that remain navigable after the shift are determined by what has been translated, preserved, or actively maintained across the discontinuity.

The displacement of minority languages by dominant ones—French displacing regional languages in metropolitan France, English reorganizing the linguistic landscape of colonial territories, Spanish restructuring prior Mesoamerican semantic systems—is precisely a restructuring of the reachability graph. Prior states (oral traditions, untranslated texts, contextual knowledge embedded in grammatical structure) are not destroyed in the moment of displacement. They become unreachable by the paths available to subsequent generations, who inherit the new navigational infrastructure.

The persistence of Arabic as a living navigational system across fourteen centuries of geographical dispersal and political transformation is an instructive counterexample. Arabic maintains its reachability through continuous traversal: liturgical practice, Quranic recitation, classical literary education, and an unbroken tradition of commentary and cross-referencing. The classical texts are not merely preserved; they are *continuously reached*, repeatedly, by a large population of agents who maintain the transformations that make them accessible. This is active path maintenance rather than archival storage.

The contrast between the two cases—a language displaced without maintained paths versus a language kept reachable through continuous practice— illuminates the constructive requirement. It is not enough to record a language; the navigational infrastructure must also be sustained.

The same logic applies to institutional language policy. Quebec's *Charte de la langue* *française* (1977) is typically analyzed as a policy of linguistic protection or identity assertion. From the present perspective, it is better understood as a constraint on reachability: it enforces, through legal means, the maintenance of edges in the francophone semantic graph that would otherwise be eliminated by the dominant pressure of English-language commercial and administrative infrastructure. The policy does not restore already-lost paths; it acts as a boundary condition that prevents certain transitions from occurring, maintaining connectivity to prior semantic states that would otherwise become unreachable within a generation.

This is not an analogy to the technical case. It is the same mechanism operating through different implementation media. The reachability graph of a language is damaged by the same structural operations—edge deletion, path compression, audience redirection—that damage the reachability graphs of digital knowledge systems. The timescales differ; the underlying structure does not.

5 The Constructive Requirement

The analysis above yields a requirement, not merely a diagnosis. If the goal of a memory system is semantic continuity—the ability of future agents to reach past meaning through available transformations—then the design question shifts from "how do we store information?" to "how do we ensure that future states can reach past states through valid paths?"

This is a stronger condition than preservation. It requires attending to three things simultaneously:

1. **Node persistence**: states must continue to exist.
2. **Edge maintenance**: the transformations that connect states must remain available and traversable.
3. **Path coherence**: the composition of available transformations must remain capable of connecting present states to past ones.

Current systems are organized primarily around the first condition. Archival standards, redundant storage, and persistent identifiers address node persistence. The second and third conditions are largely left to market incentives (search engine indexing), institutional inertia (citation practices), and individual effort (translation, commentary, cross-referencing).

The asymmetry is not incidental. Node persistence is a technical property that can be specified, measured, and contractually guaranteed. Edge maintenance and path coherence are relational properties that require ongoing coordination among the agents who constitute the system. They cannot be archived; they must be continuously performed.

This points toward a broader framework in which meaning is treated not as a stored object but as a property of a system under transformation—something that must remain *invariant under the transformations that the system makes available to its agents*. Constraint-based accounts of meaning, in which semantic content is defined by what it rules out rather than what it encodes, capture this naturally: to preserve meaning is to preserve the constraints that make certain transitions valid and others not. When those constraints are lost—when the paths are severed—the content may persist as a physical inscription while losing its operative semantic character.

This connection points beyond the scope of the present paper toward richer field-theoretic treatments of meaning under constraint, where the reachability condition developed here finds a more formal elaboration. What the present paper establishes is the prior step: that reachability, not preservation, is the invariant that matters for semantic continuity, and that modern information systems are systematically organized to maintain the wrong thing.

6 Conclusion

We have argued that the canonical framework for thinking about knowledge systems—preservation—conflates two distinct properties that can fail independently. A state can be preserved while becoming unreachable. A path can exist while being gated, corrupted, or severed. A system can maintain all its components while eliminating the navigation structure that makes those components useful.

The five failure modes examined here—namespace laundering, encyclopedic consolidation, access restriction, link rot, and archival isolation—are not pathological edge cases. They are the normal operating conditions of contemporary digital infrastructure, each representing a different mechanism by which reachability decays while preservation is maintained. The sixth case, the dense local system, shows that the inverse is also possible: reachability without archival infrastructure, sustained entirely by path redundancy.

The extension to natural language systems shows that this is not a problem of digital technology but a structural feature of any system that separates the storage of semantic states from the maintenance of the paths among them. The timescale changes; the structure does not.

The constructive implication is that continuity requires active path maintenance—continuous traversal, cross-referencing, translation, and institutional constraint—not merely storage. What must be preserved is not the object but the relation: the navigable connection between what has been said and what can still be reached.

References

[1] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423.

[2] Floridi, L. (2011). *The Philosophy of Information.* Oxford University Press.

[3] Paskin, N. (2010). Digital object identifier (DOI®) system. In *Encyclopedia of Library and Information Sciences*, 3rd ed. Taylor & Francis.

[4] Klein, M., Van de Sompel, H., Sanderson, R., Shankar, H., Balakireva, L., Liu, K., & Moretto, R. (2014). Scholarly context not found: One in five articles suffers from reference rot. *PLOS ONE*, 9(12), e115253.