

Semantic Fidelity Project Lexicon

A Working Vocabulary for Meaning Preservation Across AI and Representational Systems

A. Jacobs - Semantic Fidelity Project

Purpose

AI systems transform meaning through retrieval, compression, generation, memory, and recursive reuse. These transformations can preserve factual fragments while changing emphasis, implication, context, hierarchy, or intent. The output remains coherent, yet the meaning that organized the source becomes less recoverable.

This lexicon provides a compact vocabulary for describing that loss. It connects semantic failures inside AI workflows to the broader representational conditions that produce them. The terms are diagnostic rather than exhaustive. They are designed to clarify where meaning changes, why the change survives, and which form of corrective contact has weakened.

Semantic fidelity is not the preservation of words. It is the preservation of the relationships that make those words mean what they mean.

How the Vocabulary is Organized

Conditions describe what is being preserved or lost: semantic fidelity, semantic continuity, semantic drift, and reality drift.

Mechanisms explain how loss accumulates: recursive mediation, recursive compression, context loss, and constraint collapse.

AI Failures identify where workflows distort meaning: retrieval drift, interpretation failure, memory drift, and agent drift.

Optimization explains why degradation can look successful: proxy substitution, evaluation blindness, and synthetic coherence.

Ecosystems describe how outputs reshape future inputs: recursive synthetic feedback, semantic entropy, and provenance failure.

SECTION I - Foundational Conditions

The project begins with the difference between coherence and preserved contact.

Semantic Fidelity

The degree to which meaning, intent, context, and interpretive structure survive retrieval, transformation, compression, generation, and reuse.

Structural pattern. A representation preserves the relationships that made the source meaningful.

Distinction. Factual correctness asks whether a claim is true. Semantic fidelity asks whether the transformed representation still carries the same meaning.

Semantic continuity

The preservation of meaningful alignment across time and across multiple transformations. Continuity allows later representations to remain intelligible in relation to their originating conditions.

Structural pattern. Meaning remains traceable across a chain of transformations.

Distinction. Fidelity can describe one transformation. Continuity describes preservation across the whole chain.

Semantic drift

The gradual mutation of meaning while surface coherence or factual fragments remain intact. Drift may alter framing, implication, tone, hierarchy, or purpose without producing an obvious error.

Structural pattern. Meaning changes while the representation still appears valid.

Distinction. Semantic drift concerns meaning. Representation drift concerns the representation's relationship to its source or referent.

Reality Drift

The broader condition in which a system remains operational and internally coherent while progressively losing corrective contact with the reality it was built to represent or act upon.

Structural pattern. Operational continuity persists while external alignment weakens.

Distinction. Reality drift names the system-level condition. Semantic drift is one mechanism by which it can develop.

SECTION II - Representation and Reference

Meaning depends on relationships among representation, source, context, and use.

Representation drift

The progressive divergence of a representation from the condition, context, or meaning it originally encoded while the representation continues circulating and influencing downstream systems.

Structural pattern. The representation remains usable after reference integrity weakens.

Distinction. Semantic drift changes meaning. Representation drift weakens correspondence between representation and referent.

Reference integrity

The degree to which a representation remains connected to an identifiable source, object, event, or originating condition and can still be corrected by it.

Structural pattern. The referent remains available as a constraint on interpretation.

Distinction. Provenance records lineage. Reference integrity describes whether that lineage still supports meaningful correction.

Provenance integrity

The preservation of source identity, transformation history, and evidentiary lineage across a representational chain.

Structural pattern. A later output can be traced through the operations that produced it.

Distinction. Traceability alone does not guarantee fidelity, but fidelity becomes difficult to test without traceability.

Provenance failure

The condition in which the origin or transformation history of a representation cannot be reliably reconstructed. Source and synthesis become difficult to separate.

Structural pattern. A representation circulates without recoverable lineage.

Distinction. Provenance failure concerns origin. Semantic drift concerns meaning. They often compound each other.

SECTION III - Mechanisms of Loss

Most degradation accumulates through ordinary transformations rather than a single dramatic error.

Recursive mediation

A condition in which reality is increasingly encountered through representations of prior representations: summaries, embeddings, interfaces, feeds, model outputs, and institutional narratives.

Structural pattern. Representations become inputs for later representations.

Distinction. Mediation describes the layered condition. Recursion describes its reuse as new input.

Recursive Compression

The repeated reduction of semantic complexity as compressed representations become inputs for further compression. Each pass retains what is easiest to transmit and weakens what depends on context or relational structure.

Structural pattern. Compression compounds because earlier omissions become the next layer's starting reality.

Distinction. Compression is not automatically loss. Recursive compression becomes dangerous when omitted structure cannot re-enter.

Context loss

The removal or weakening of conditions needed to interpret a representation correctly, including scope, audience, time, purpose, uncertainty, and surrounding relationships.

Structural pattern. Content survives after the conditions governing its meaning disappear.

Distinction. Context drift describes changing context across a workflow. Context loss names context that is no longer available.

Constraint Collapse

The weakening of material, contextual, or procedural constraints that once prevented a representation from drifting too far from reality or intent.

Structural pattern. Internal coherence replaces external resistance.

Distinction. Constraint collapse removes correction. Semantic drift describes what can then happen to meaning.

Meaning debt

The accumulated burden created by locally convenient omissions, simplifications, and unresolved ambiguities. Each shortcut appears manageable until later systems depend on the compressed result.

Structural pattern. Small semantic losses become costly downstream dependencies.

Distinction. Meaning debt emphasizes the future cost created by accumulated semantic loss.

SECTION IV - Retrieval and Interpretation

Relevant information can be present while the system still reconstructs the wrong meaning.

Retrieval drift

The gradual misalignment between what a query or task requires and what a retrieval system repeatedly surfaces. Retrieved material may remain topically adjacent while becoming structurally wrong for the intended interpretation.

Structural pattern. Relevance signals substitute for task-specific fit.

Distinction. Retrieval failure returns missing or irrelevant material. Retrieval drift can return plausible material that redirects the frame.

Embedding substitution

The treatment of vector proximity as if it were semantic equivalence. Statistical adjacency becomes a substitute for shared role, implication, or conceptual structure.

Structural pattern. Similarity is accepted as sameness.

Distinction. Embeddings are representations. The failure occurs when their compressed relation is treated as the full meaning relation.

Interpretation Failure

The incorrect integration, framing, or use of information that was successfully retrieved. The necessary evidence may be present while the relationships among the evidence are reconstructed incorrectly.

Structural pattern. Correct context is assembled into the wrong meaning.

Distinction. Retrieval quality concerns what enters the context. Interpretation quality concerns what the system does with it.

Context drift

The progressive mutation or disappearance of contextual constraints across long conversations, iterative prompts, handoffs, or multi-step workflows.

Structural pattern. Later steps inherit a reconstructed version of the original conditions.

Distinction. Context loss may happen once. Context drift accumulates across a sequence.

Retrieval success does not imply interpretation success. Presence is not preservation.

SECTION V - Memory and Agency

Persistence introduces a new problem: later actions inherit transformed versions of earlier meaning.

Memory drift

The gradual divergence of stored summaries, preferences, state, or user models from the conditions that originally made them valid. Each update can reinterpret rather than merely preserve prior state.

Structural pattern. Persistent memory recursively rewrites its own context.

Distinction. Context drift occurs within an active chain. Memory drift persists across interactions and time.

Agent drift

The progressive divergence of an agent's actions from the originating objective across planning, retrieval, tool use, memory updates, delegation, and feedback cycles.

Structural pattern. Locally reasonable steps accumulate into global misalignment.

Distinction. Task failure may be obvious. Agent drift can produce successful-looking completion of the wrong objective.

Instruction sedimentation

The accumulation of prior prompts, inferred preferences, policies, summaries, and tool outputs into an implicit operating context that later actions treat as authoritative.

Structural pattern. Past representations harden into present constraints.

Distinction. Memory drift changes stored meaning. Instruction sedimentation describes how stored layers acquire governing force.

Procedural completion

The successful execution of a specified workflow without resolution of the underlying need. Completion is measured by performed steps rather than restored contact with the goal.

Structural pattern. The procedure closes while the problem remains open.

Distinction. Procedural completion is a common operational form of proxy substitution.

SECTION VI - Optimization and Evaluation

Systems often become better at producing the evidence by which they are judged.

Proxy substitution

The replacement of an underlying objective by a metric, benchmark, reward signal, procedure, or other measurable representation of that objective.

Structural pattern. The indicator becomes the operational definition of success.

Distinction. Proxy optimization improves the stand-in. Proxy substitution occurs when the stand-in is accepted as the thing itself.

Representational optimization

The optimization of content or evidence for circulation, ranking, retrieval, compression, compliance, or machine legibility rather than for preserved reference and meaning.

Structural pattern. The representation adapts to the system that processes it.

Distinction. Optimization changes form. Drift begins when those changes weaken fidelity and cannot be corrected.

Evaluation blindness

The inability of a benchmark or evaluation process to detect semantic degradation that falls outside its measurable criteria.

Structural pattern. Local scores confirm success while omitted relationships deteriorate.

Distinction. A poor evaluation may be inaccurate. Evaluation blindness can be internally reliable and still measure too little.

Synthetic Coherence

The appearance of fluent and mutually reinforcing meaning produced by representations that no longer have sufficient grounding or independent support.

Structural pattern. Coherence detaches from contact.

Distinction. Hallucination is usually framed as a false statement. Synthetic coherence describes a larger structure that feels valid because its parts fit together.

Survivable Wrongness

Misalignment that remains usable, plausible, or institutionally compatible enough to persist without triggering correction.

Structural pattern. The system continues, so the error becomes part of normal operation.

Distinction. Failure is an event. Survivable wrongness is a condition that can reproduce itself.

SECTION VII - Recursive Ecosystems

Generated representations increasingly shape the environments from which future systems learn and retrieve.

Recursive synthetic feedback

The process by which generated outputs become future inputs for training, retrieval, ranking, memory, policy, and further generation.

Structural pattern. Synthetic representations alter the conditions of later representation.

Distinction. Feedback is not inherently degrading. Risk rises when synthetic material loses provenance or displaces independent sources.

Semantic entropy

The increasing disorder and instability of meaning across a representational environment. Distinctions become harder to preserve, origins harder to recover, and interpretations less reliably reconstructable.

Structural pattern. Repeated transformation increases uncertainty about meaning and source.

Distinction. Semantic drift may have direction. Semantic entropy describes declining order across the wider environment.

Lexical decay

The narrowing of expressive range as precise, rare, situated, or metaphorically rich language is replaced by generic and statistically safe formulations.

Structural pattern. Fluency persists while semantic texture contracts.

Distinction. Lexical decay is visible at the level of language. Ground erosion describes the broader loss of background structure.

Ground erosion

The weakening of tacit background, lived context, hierarchy, absence, and material reference that gives explicit language its weight.

Structural pattern. Foreground content remains while the background that organizes significance disappears.

Distinction. Context loss removes stated conditions. Ground erosion weakens the larger field from which interpretation draws.

SECTION VIII - Evaluation Vocabulary

Semantic fidelity requires visibility into transformations, not only judgments about final outputs.

Semantic observability

The capacity to inspect how meaning changes across retrieval, transformation, memory, generation, and reuse. Observability makes the semantic chain available for diagnosis.

Structural pattern. Transformations can be compared rather than inferred from the final output alone.

Distinction. Provenance shows where an output came from. Semantic observability examines what changed along the way.

Fidelity evaluation

The assessment of whether a transformation preserved the relevant intent, relationships, context, uncertainty, and reference structure for its use.

Structural pattern. Evaluation is defined by what meaning must survive, not only by textual similarity.

Distinction. Accuracy is one component. Fidelity evaluation also tests framing, omission, hierarchy, and downstream use.

Constraint retention

The preservation of requirements, boundaries, exceptions, and real-world conditions across a workflow. Retained constraints continue limiting what counts as a valid output.

Structural pattern. Later steps remain answerable to originating conditions.

Distinction. Context supplies meaning. Constraints determine what transformations remain acceptable.

Corrective Contact

A system's practical ability to encounter evidence, consequences, cases, or judgment capable of contradicting its internal representations.

Structural pattern. External reality can still interrupt internal coherence.

Distinction. Feedback is any returned signal. Corrective contact requires a signal that remains meaningfully connected to the underlying condition.

How the Terms Fit Together

The vocabulary follows a six-stage sequence through which meaning is represented, transformed, optimized, and reused.

1. Source - Reality, intent, or experience contains more structure than any representation can preserve.

2. Representation - Data, language, metrics, and evidence preserve some relationships while making others less visible.

3. Transformation - Retrieval, ranking, summarization, and generation can alter meaning through context loss, interpretation failure, and weakened constraints.

4. Optimization - Rewards, benchmarks, compliance requirements, and circulation pressures can make distorted representations appear successful.

5. Recursion - Representations reused as memory, evidence, or future input carry previous losses forward and compound them.

6. System State - The system remains coherent and operational while semantic drift gradually widens into reality drift.

Working Distinctions

Semantic fidelity describes preservation. Semantic drift describes change in meaning. Representation drift describes divergence from the referent. Reality drift describes the system-level condition that emerges when weakening representations continue to organize action.

The remaining terms explain how that condition develops and survives. Recursive compression compounds what has already been lost. Proxy substitution gives the compressed representation operational authority. Evaluation blindness prevents the loss from being detected, while provenance failure makes the original meaning difficult to reconstruct.

The absence of visible failure is not proof that meaning has been preserved. The central principle of the project is that coherence is not contact. A representation can remain clear, useful, and operational after it has stopped preserving the relationships that made it meaningful.

This working lexicon defines terms by their diagnostic role. Definitions may be refined as the Semantic Fidelity Project develops.

