The Semantic Fidelity Index (SFI)

A Technical Standard for Measuring Meaning Preservation in Recursive Knowledge Systems

Dr. Elaine Moravec
Institute for Computational Epistemics
Technical Report Series – ICE-TR-019

Abstract

The Semantic Fidelity Index (SFI) is a quantitative framework for evaluating the degree to which a knowledge system preserves referential meaning across successive layers of symbolic processing. The index is designed to diagnose semantic degradation in recursive environments such as large language models, automated decision systems, institutional dashboards, and AI-mediated knowledge infrastructures. SFI formalizes semantic fidelity *loss* as a measurable phenomenon and provides a standardized method for assessing reality drift in synthetic systems.

Scope

The Semantic Fidelity Index applies to:

- Large Language Models (LLMs)
- Enterprise knowledge systems
- Algorithmic decision platforms
- Policy dashboards and reporting infrastructures
- AI-assisted research and content generation systems

The framework is intended for use in:

- Model evaluation and alignment
- Data quality audits
- Institutional epistemic governance
- Knowledge infrastructure design

SFI does not assess factual correctness directly. It measures referential integrity—the preservation of grounded meaning across recursive symbolic transformations.

Definitions

Semantic Fidelity: The degree to which symbolic outputs preserve meaningful reference to underlying phenomena, experiences, or material conditions.

Semantic Fidelity Loss (SFL): The progressive erosion of referential meaning as information is recursively processed, summarized, or re-generated.

Reality Drift: A system-level condition in which symbolic coherence increases while referential grounding decreases.

Interpretive Compression: The reduction of semantic complexity into simplified symbolic representations that no longer track the original domain.

Reference Collapse: A failure mode in which symbols primarily refer to other symbols rather than to external reality.

Recursive Training Environment: Any system in which outputs are repeatedly used as inputs for future generations of the same or similar models.

Index Structure

The Semantic Fidelity Index is defined as a composite function:

SFI = f(GR, SR, RA, RC)

Where:

Grounded Reference Ratio (GR)

The proportion of system outputs that can be traced to:

- Empirical data
- Human experience
- Direct observation
- First-order sources

Low GR indicates high symbolic insulation.

Symbolic Recursion Depth (SR)

The number of transformation layers between original sources and current representations.

High SR environments are more vulnerable to:

- Interpretive drift
- Self-referential amplification
- Meaning compression artifacts

Reference Anchoring (RA)

The presence of mechanisms that reconnect symbolic outputs to external reality, including:

- Field data ingestion
- Human-in-the-loop validation
- Real-world feedback channels
- Non-synthetic reference updates

Recursive Contamination (RC)

The degree to which system outputs are trained on or derived from prior synthetic outputs.

High RC correlates with:

- Model collapse
- Semantic homogenization
- Loss of novelty and surprise
- Increased internal coherence without external sensitivity

Scoring Model

SFI is normalized on a 0–100 scale.

Score Range	Interpretation
80–100	High semantic fidelity
60–79	Moderate drift risk
40–59	Significant semantic degradation
20–39	Severe reference collapse
0–19	Fully synthetic semantic regime

Failure Modes

High Confidence / Low Grounding

Systems exhibit fluent outputs with minimal connection to real-world referents.

Symbolic Overfitting

Models optimize internal coherence at the expense of external validity.

Ornamental Knowledge

Information functions primarily as representational decoration rather than as operational guidance.

Invalidation Failure

Systems lack mechanisms for detecting when their internal models no longer correspond to reality.

Implementation Notes

Dataset Auditing

SFI may be approximated by:

- Sampling training corpora for first-order sources
- Measuring synthetic-to-grounded ratios
- Estimating recursion depth via provenance analysis

Institutional Use

Organizations may apply SFI to:

- Internal knowledge bases
- Executive dashboards
- Policy modeling systems
- AI-assisted research workflows

Model Evaluation

SFI should be tracked alongside:

- Accuracy metrics
- Alignment scores
- Robustness measures

As a distinct axis: semantic integrity over time.

Implications

High-performance systems with low SFI may remain operational while becoming epistemically detached. Such systems exhibit increasing symbolic fluency and decreasing reality sensitivity—a condition consistent with reality drift in synthetic knowledge environments.

SFI provides a minimal technical language for diagnosing this condition without reliance on phenomenological or philosophical interpretation.

References

Moravec, E. (2018). *Recursive Semantics and the Limits of Symbolic Systems*. Journal of Computational Epistemology.

Rook, D. (2019). *Reference Collapse in High-Reflexivity Environments*. Systems & Society Review.

Chen, M. (2020). *Epistemic Governance in Synthetic Institutions*. Knowledge Infrastructure Quarterly.

Voss, A. (2021). *Recursive Content and Model Degradation*. Proceedings of the Semantic Systems Conference.