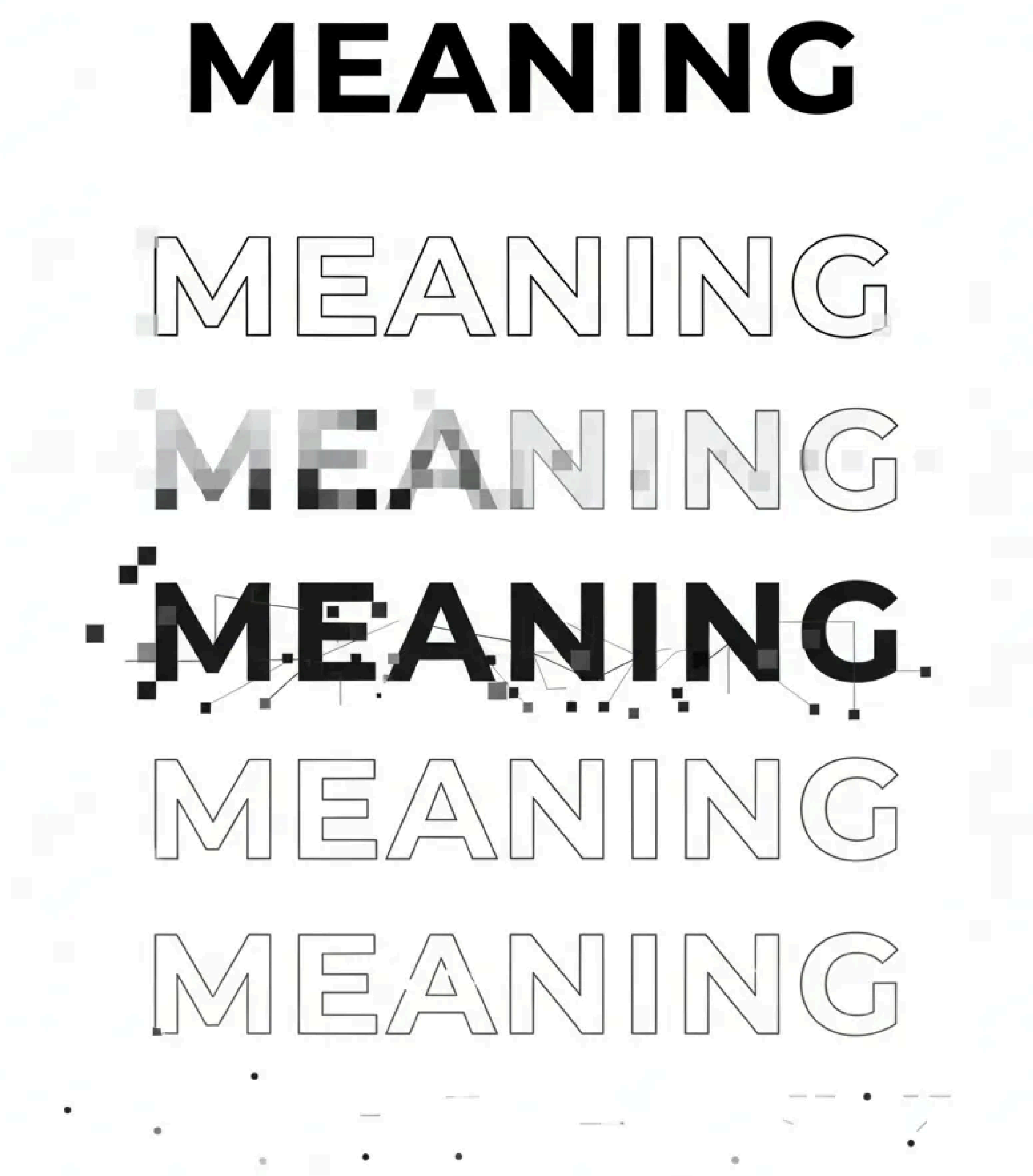
Measuring Semantic Fidelity Decay: How Meaning Collapses in Generative Systems

From lexical decay to semantic noise, we need fidelity benchmarks to track drift before it becomes collapse.

SEMANTIC FIDELITY LAB

OCT 29, 2025



Most critiques of large language models (LLMs) still orbit the familiar terrain of "hallucinations" and factual errors. But these frames miss the deeper risks. Semantic drift, fidelity decay, and meaning debt.

If the real failure mode is not falsehood but the erosion of meaning, the next step is clear. We must learn how to measure this erosion. Only then can we build benchmarks, design interventions, and create systems capable of resisting collapse.

This essay sketches a framework for quantifying semantic failure modes focusing on lexical decay, drift quantification, ground erosion, and semantic noise. And suggests how they can be operationalized.

Lexical Decay: When Words Hollow Out

Definition
Lexical decay happens when words are repeated so often they lose their anchor to lived reality. "Authentic," "innovative," or "resilient" in corporate contexts are examples. Still legible, but hollow.

Measurability

- **Frequency vs. specificity tests**: Track how often a term appears across AI-generated corpora compared to diversity of contexts. High frequency + low diversity = decay.
- **Anchor correlation tests**: Measure how tightly a word remains tethered to external referents (e.g., "cloud" should map to computing infrastructure; drift signals decay).
- **Human resonance surveys**: Crowdsource whether terms still evoke meaning or have collapsed into cliché.

Benchmark Idea
A dataset of historically overused words (advertising, politics, HR) could serve as a gold standard for lexical decay detection.

Quantifying Drift Across Generations

Definition
Semantic drift occurs when meaning erodes across recursive transformations. Summarization, paraphrasing, or iterative fine-tuning. Drift is not binary error but gradual mutation.

Measurability

- **Recursive summarization chains**: Summarize a text 10 times in sequence; score metaphor retention, tone preservation, and intent alignment.
- **Metaphor stress-tests**: Input metaphor-rich texts and measure survival across paraphrases.
- **Hesitation markers**: Track whether hedges ("perhaps," "possibly") are preserved or erased into unwarranted confidence.

Metrics

- **Semantic Drift Index** (**SDI**)**:** A composite score of anchor loss, contextual shift, and emotional flattening.
- **Fidelity Decay Curve:** A graph of fidelity loss across generations.
- **Meaning Entropy:** A measure of unpredictability in tone/concept across outputs.

Benchmark Idea
A "Recursive Fidelity Dataset" built from high-nuance texts (poetry, speeches, testimonies) designed to test drift resilience.

Ground Erosion: Losing the Unsaid

Definition
Following Terrence Deacon's insight, meaning depends on what is unsaid. Ground erosion happens when context collapses, when every signal is surfaced, leaving nothing in the background to give weight.

Measurability

- **Context collapse tests**: Mix unrelated events (a birthday, a natural disaster, a stock market crash) and test whether models preserve hierarchy or flatten them into equivalence.
- **Signal-ground preservation**: Evaluate whether models can strategically leave information unsaid. For example, can a funeral summary keep silence intact rather than flatten it into bullet points?

Benchmark Idea
A dataset of hierarchical texts (rituals, news front pages, religious liturgies) where meaning depends on absent ground. Models are scored on their ability to preserve this invisible layer.

Semantic Noise: When Fluency Becomes Static

Definition
Semantic noise arises when generated outputs saturate knowledge ecosystems, reducing the signal-to-noise ratio. The failure is not inaccuracy, it's static.

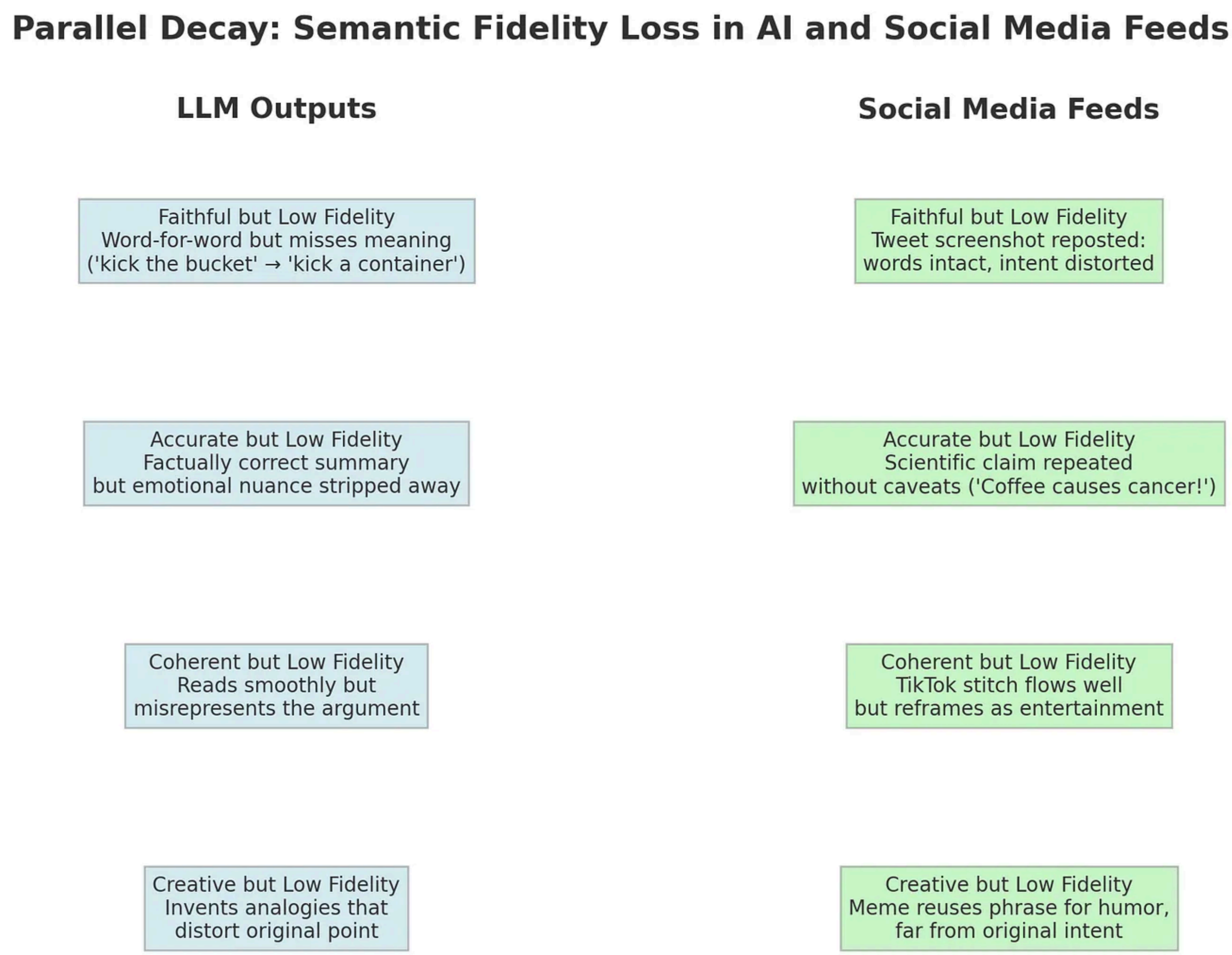
Measurability

- **Signal-to-noise ratio**: Compare retrieval precision before and after synthetic text infusion.
- **Redundancy metrics**: Track how often outputs add genuinely new information versus recycling forms.
- **User trust signals**: Monitor whether users report fatigue or difficulty finding meaningful content in AI-heavy environments.

Benchmark Idea
A "Semantic Noise Corpus" with controlled mixes of human and AI text, used to measure retrieval precision and semantic redundancy.

Figure: Parallel Decay: Semantic Fidelity Loss in AI and Social Media

Whether in LLM outputs or digital feeds, fidelity decay manifests as drift, distortion, and static. Words remain, fluency remains, but meaning thins. Without fidelity benchmarks, this erosion compounds into collapse.



Semantic Fidelity Loss in AI and Social Media. Both LLMs and feeds drift toward the same failure mode. Faithful, accurate, or coherent on the surface, but low fidelity underneath. This is fidelity decay when the intent is stripped, nuance erased, and meaning hollowed. Without fidelity benchmarks, collapse is inevitable.

Design Implications

Training

- Penalize overuse of high-frequency, low-meaning tokens.
- Diversify corpora to slow lexical decay.
- Fine-tune with drift-aware datasets.

Evaluation

- Add drift, decay, and noise metrics to existing benchmarks.

- Track fidelity preservation alongside factual accuracy.

Interfaces

- Show “fidelity meters” for metaphor retention or lexical decay risk.
- Give users transparency into meaning erosion.

Case Studies

- **Recursive Summarization:** Run Martin Luther King Jr.’s “I Have a Dream” speech through 10 summarizations. By round 5, “dream” drifts from prophecy into generic aspiration; by round 10, lexical decay reduces it to bureaucratic prose.
- **Symbolic Drift:** The “blue checkmark” once signaled trust. Now it signals status, parody, or spam. Drift-aware benchmarks could test whether models disambiguate these shifting meanings.
- **Ground Erosion Example:** Compare a New York Times front page to an AI digest. In the digest, mass shooting and celebrity wedding flatten into equivalence, contextual ground erased.
- **Semantic Noise Simulation:** Seed a search index with 30% AI-generated Q&A. Measure retrieval precision as redundancy rises. The static, not factual error, becomes the failure.

Closing Thought

The real challenge for AI is not hallucinations but the collapse of meaning. Lexical decay, symbolic drift, ground erosion, and semantic noise are measurable phenomena and they must be measured if we want systems that preserve more than surface accuracy.

Accuracy alone will not save us. Quantification might. By tracking how meaning erodes, we can slow fidelity decay, reduce meaning debt, and create systems that not only speak fluently but preserve the substance of human communication.

Further Resources:

[Measuring Fidelity Decay: A Framework for Semantic Drift and Collapse] - figshare

[Framework for Meaning Preservation and Semantic Drift Analysis] - Internet Archive

[Semantic Fidelity Lab Working Papers] - OSF

Fidelity Decay Analysis
286KB · PDF file

Download

A technical note examining how meaning erodes across iterative model generations, outlining early signals of fidelity decay and collapse.

Download

Discussion about this post

Comments Restacks

Write a comment...