

Day 6 — Semantic Validation (Claim \leftrightarrow Context Entailment)

Goal

Ensure every factual claim in the answer is supported by approved context.

Claim Extraction

Answers are decomposed into atomic factual claims.

Opinions and hedging language are ignored.

Entailment

Each claim is checked against context using an LLM as a verifier.

Only ENTAILED is acceptable.

Fail-Closed Design

No context \rightarrow NOT_ENTAILED.

UNKNOWN \rightarrow failure.

Any single failure invalidates the answer.

Outcome

This creates a semantic firewall against hallucinations.