

Layer

Data

Purpose

Monitor data quality, freshness, and lineage

Model

Track model performance and version health

Behavior

Observe agent and workflow execution patterns

Semantic

Evaluate meaning, reasoning, and policy alignment

Key telemetry types

Schema validation, freshness timestamps, lineage records, drift detection

Accuracy, latency, token usage, version metadata, embedding quality

Tool call sequences, decision paths, escalation events, error patterns

Groundedness scores, intent alignment, policy tags, faithfulness scores, reasoning traces

Tooling examples

Data catalogs, lineage tools, quality monitors

MLOps platforms, model registries, performance dashboards

Tracing systems, agent observability platforms, workflow monitors

Evaluation frameworks, LLM-as-judge systems, governance platforms