



Tracing the Agent's Mind

9-11 Sep. 2026

IEEE ASYU 2026



Keynote Speaker:

Mustafa Dayioğlu, Ph.D.

TÜBİTAK Senior Chief Researcher

Abstract

Production AI agents make thousands of tool-calling decisions daily, yet observability stops at the model boundary. OpenTelemetry's GenAI semantic conventions capture token counts and latencies— what the LLM processed —but not why an agent selected a specific tool. Research (McKenzie et al., 2023) demonstrates inverse scaling: more capable models exhibit unpredictable tool selection patterns. This gap leaves engineers guessing during critical production failures.

We present `gen-ai-otel`, an open-source OpenTelemetry extension introducing decision-level telemetry for MCP agents. A new attribute namespace (`gen_ai.agent.*`) captures tool selection confidence, session context, permission scope validation, and baseline deviations. The zero-sidecar architecture routes telemetry through standard Collector pipelines to existing backends — Jaeger, Prometheus, or graph databases— with low overhead and cardinality-aware attributes.

A live demo reconstructs an agent's decision chain, revealing anomalies invisible to token metrics, reducing decision-debugging time.