# Edging Closer to AI Assisted Troubleshooting

## AI Assisted Dashboard Analysis

**Jing Hu — MTS 1, Software Engineer**

**Vijay Samuel — Principal MTS, Architect**

© 2026 eBay. All rights reserved. Confidential and proprietary.

# #WeAreeBay!

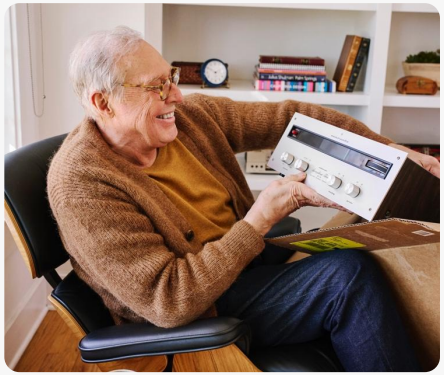
© 2026 eBay. All rights reserved. Confidential and proprietary.

# eBay Today!



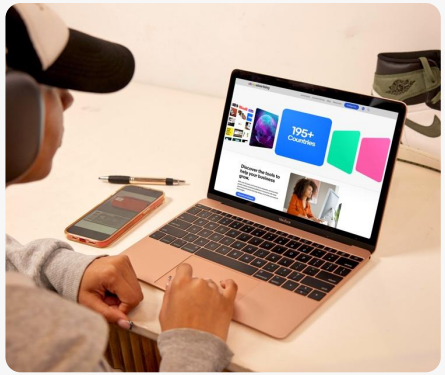

**2.8B**

Live Listings



**147M**

Total Active Buyers Worldwide




**190+**

Markets



**$79.6B**

GMV

***All data approximate as of Q3 2026, except for GMV which shows full year 2025**

© 2026 eBay. All rights reserved. Confidential and proprietary.

# 01

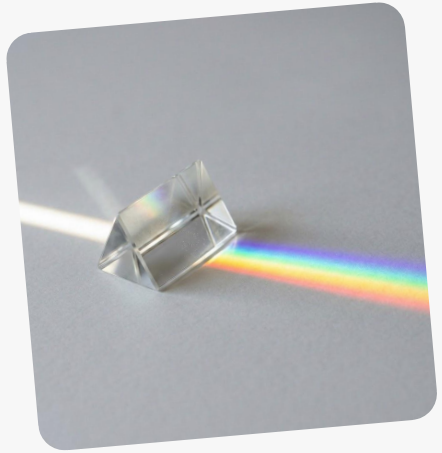
# Observability Is All About Angles...

© 2026 eBay. All rights reserved. Confidential and proprietary.

# Dashboards Are Curated Angles

- **Panels and queries:** the angles themselves
- **Variables:** the scope (which service, which region)
- **Text panels:** tribal knowledge ("if this spikes, check X")
- **offset queries:** the baseline ("compare with last week")
- **Annotations:** what changed

**The investigation plan already exists!**





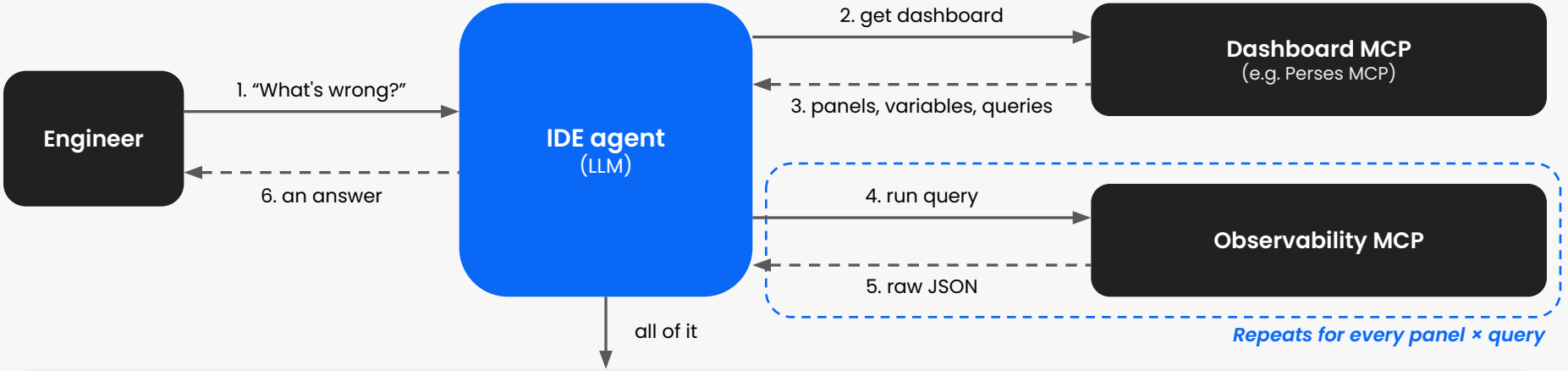



© 2026 eBay. All rights reserved. Confidential and proprietary.

# Everyone Has An MCP Server Now...

© 2026 eBay. All rights reserved. Confidential and proprietary.

# The Two-MCP Pattern




**IDE agent's context window**

| dashboard: panels, variables, queries | panel 1 · raw JSON | panel 2 · raw JSON | panel 3 · raw JSON | … | panel N · raw JSON |
|---|---|---|---|---|---|

The LLM parses, computes, compares and concludes, all from inside this window.

## Every raw result lands in the context window!

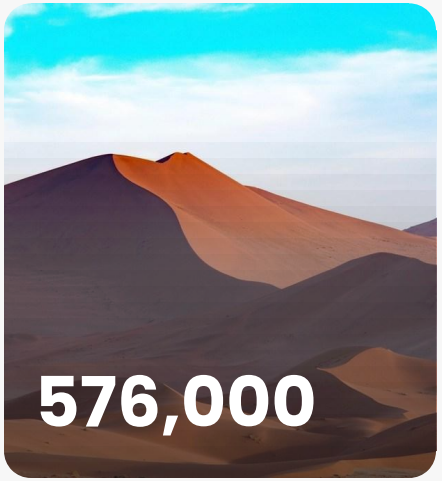
© 2026 eBay. All rights reserved. Confidential and proprietary.

# Welcome To Tokenmaxxing!

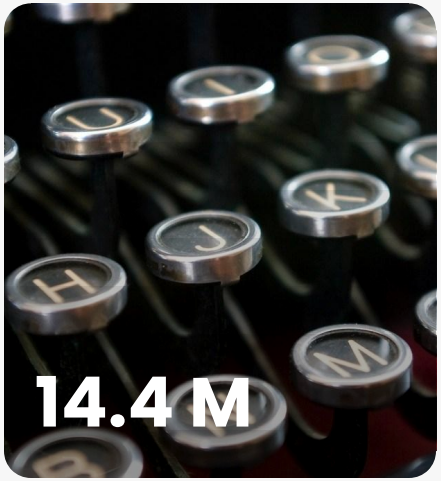
Pushing as many tokens through the model as possible, and calling it progress

© 2026 eBay. All rights reserved. Confidential and proprietary.

# Back Of The Envelope…



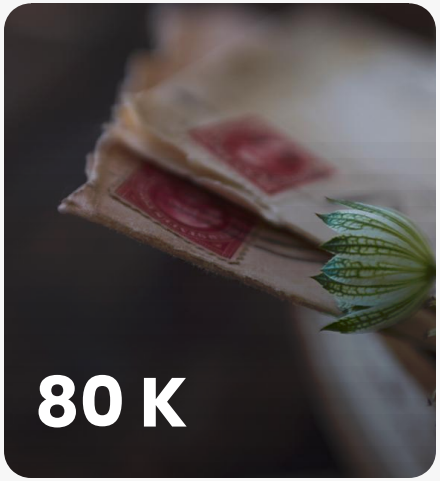

Samples: 40 panels × 20 series × 720 points (6 h at 30 s)




Characters of JSON




Tokens, and that's *optimistic*




Characters: our budget for a whole analysis step

**One panel alone is ≈ 90 K tokens!**

Illustrative; ~25 chars per sample, ~4 chars per token.

© 2026 eBay. All rights reserved. Confidential and proprietary.

# It's Not Just The Bill...

- Facts in the middle of long contexts get lost
- Tokenization alone moved arithmetic accuracy from **76% to 98%**
- 1,000 runs at temperature 0 gave **80 different answers**
- Fluent text that the data doesn't support

**On-call needs the exact opposite!**





© 2026 eBay. All rights reserved. Confidential and proprietary.

# We Fell For It Too...

© 2026 eBay. All rights reserved. Confidential and proprietary.

# Traces: Three Attempts




**Attempt 1:** raw span trees → one LLM call. HTTP 500s.




**Attempt 2:** an LLM call per trace, then a synthesis call. Flaky.




**Attempt 3:** Go computes verdicts, one LLM call narrates. Unit-tested.

**“The LLM never needs to do math over span trees.”**

Figures from our internal design notes.

© 2026 eBay. All rights reserved. Confidential and proprietary.

# 02

# Meet Dashboard Analyzer!

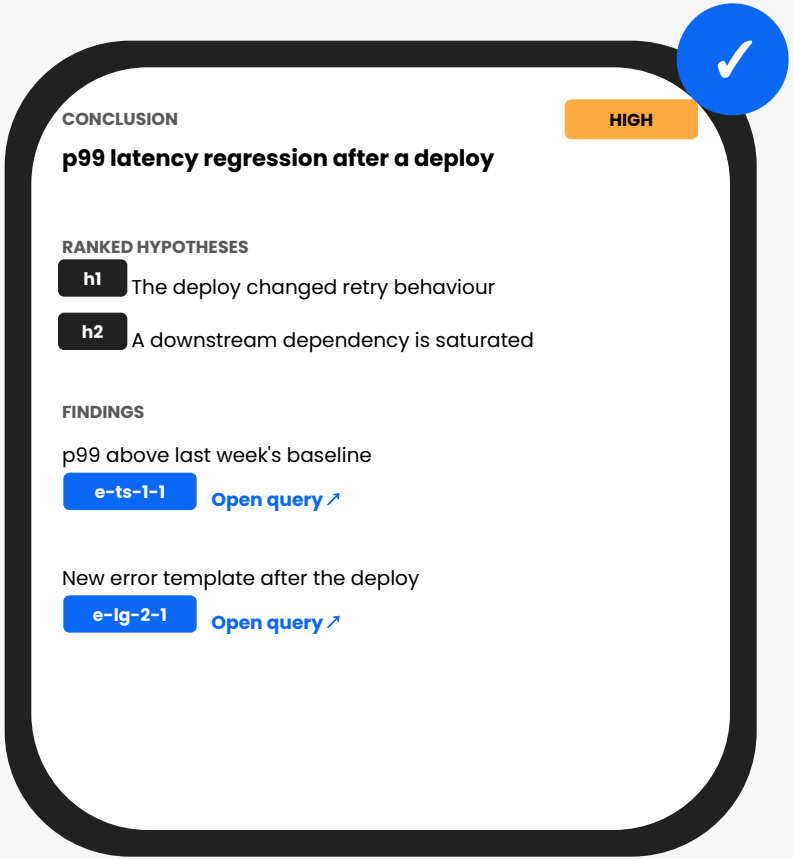
© 2026 eBay. All rights reserved. Confidential and proprietary.

# Paste A Link, Get Evidence

- **In:** a dashboard link (time range and variables included) or an ad-hoc query link
- **Out:** a conclusion with severity
- Ranked hypotheses
- Findings, each backed by evidence
- Deep links back to the exact query and time window

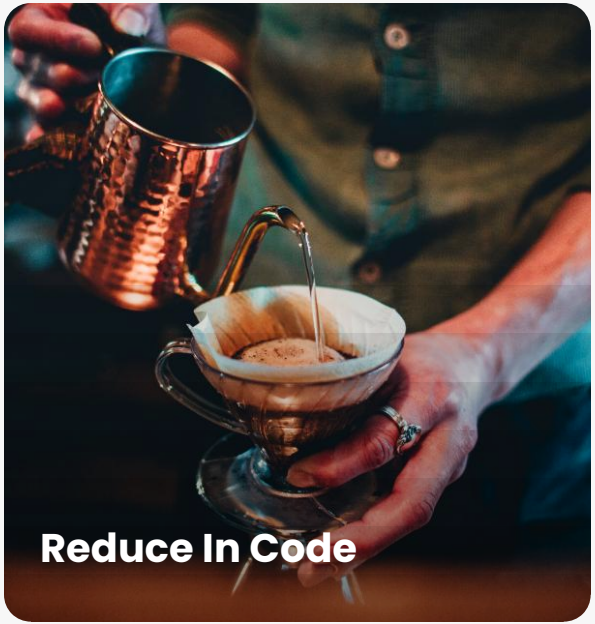
**Every claim is one click from its proof!**



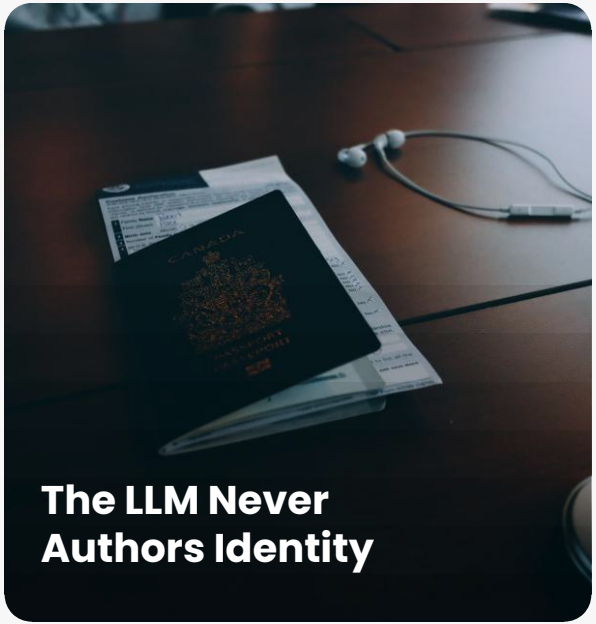



© 2026 eBay. All rights reserved. Confidential and proprietary.

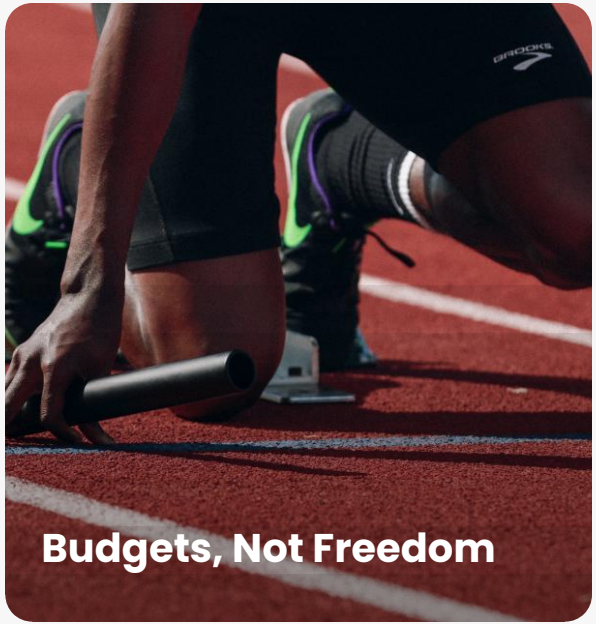
# Three Rules We Didn't Break




Parse, compute, detect and decide in Go
The model reads a small digest
The model narrates




IDs, timestamps and links come from code
The model can only refer to them




Fixed plan shape
Hard caps everywhere
No open-ended tool calling

© 2026 eBay. All rights reserved. Confidential and proprietary.

# Who Does What?

- **Code:** parse, query, statistics, patterns, verdicts
- **Code:** IDs, links, time anchoring, caps
- **LLM #1 Planner:** which panels matter
- **LLM #2 Step analyzer:** what one step's digest means
- **LLM #3 Synthesizer:** how findings connect, and what to check next

**Three LLM calls. JSON out. Zero tools.**



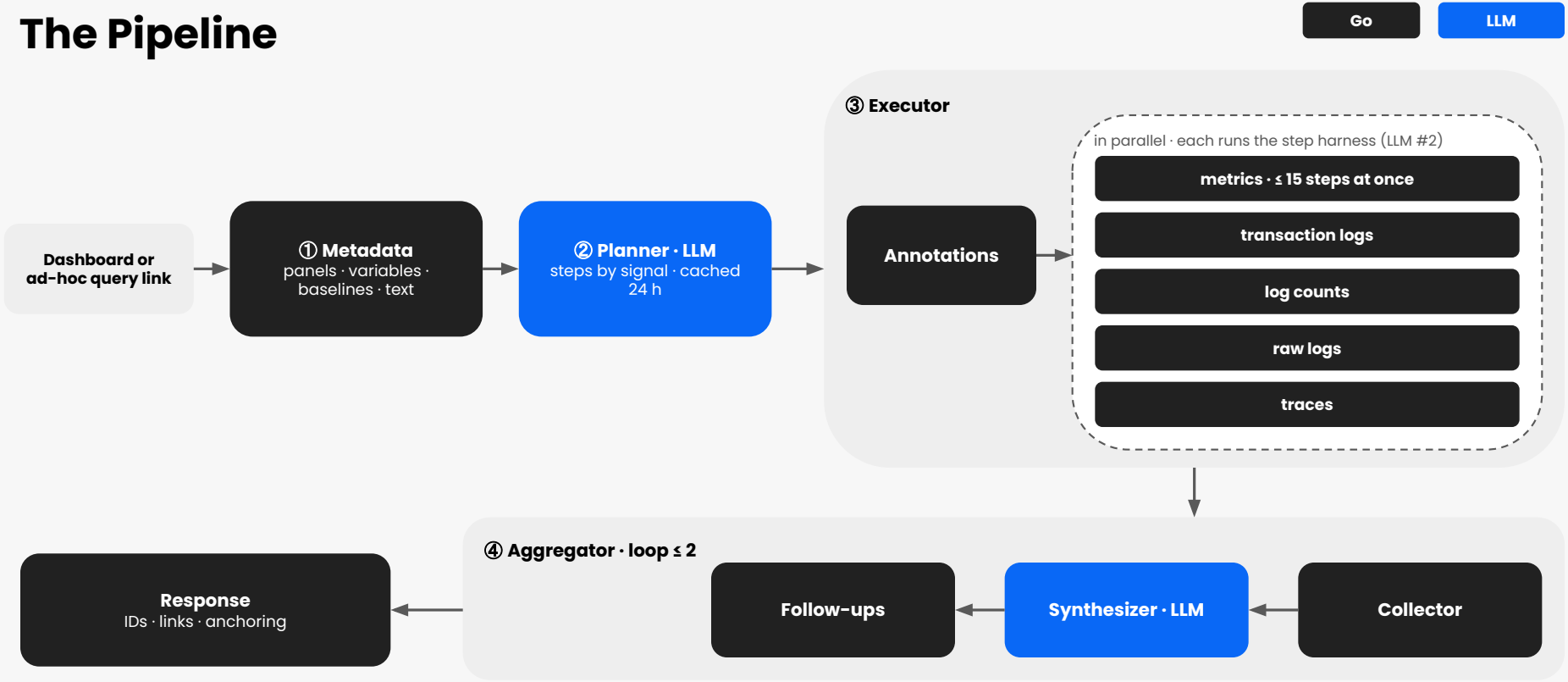


© 2026 eBay. All rights reserved. Confidential and proprietary.

# Putting It All Together…

© 2026 eBay. All rights reserved. Confidential and proprietary.

# The Pipeline



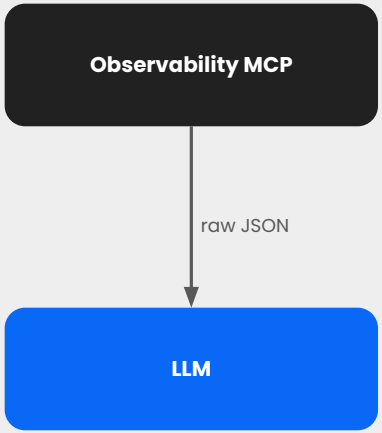

**Four stages. Three LLM calls. Everything else is Go.**

© 2026 eBay. All rights reserved. Confidential and proprietary.

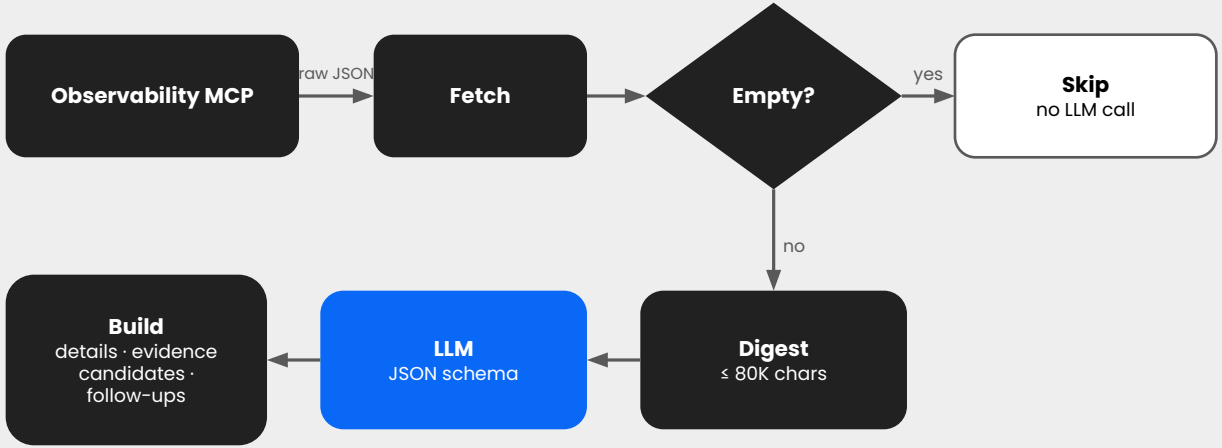
# Same MCP Servers, Different Middle

Go LLM

**Two-MCP pattern**




**One Dashboard Analyzer step**




**We kept the MCP servers. We changed what sits in between.**

© 2026 eBay. All rights reserved. Confidential and proprietary.

# Every Stage Is An Agent

- Built on **ADK for Go**, Google's open-source agent framework
- One interface: Run(ctx) → stream of events
- Agents talk through state changes on events, not function calls
- Sequential · parallel · loop, plus a **for-each** we added
- A loop stops when an agent sets Escalate

**Planner, step or whole pipeline: same shape, so we can test and compose each one.**





© 2026 eBay. All rights reserved. Confidential and proprietary.

# 03

# Where Did The Tokens Go?

© 2026 eBay. All rights reserved. Confidential and proprietary.

# Metrics: 720 Points → A Digest

- Stats in Go: min, max, mean, **P50 / P95 / P99**, trend
- Anomaly hints: analysis service first, classic checks as fallback
- Up to 10 series per query in detail; the rest summarised
- Chart shape: LTTB downsampling, **720 → 50 points**
- Baseline comparison against the offset query

**The LLM never computes a percentile!**



© 2026 eBay. All rights reserved. Confidential and proprietary.

# The Default That Would Have Lied…

- One of our MCP tools returns **5 series** by default
- On busy panels it silently drops the rest
- No error, no warning, nothing to retry
- Fix: set limits explicitly on every call (**200 / 50**)

**MCP tool defaults are part of your prompt!**



© 2026 eBay. All rights reserved. Confidential and proprietary.

# Logs: A Thousand Lines → A Handful Of Templates

- Up to **1,000 lines** per query → templates mined Drain-style, with <*> for the variable parts
- Each template: a count + up to 5 examples
- Counts and rows: median / MAD outlier test (|M| > 3.5)
- Spike **≥ 3×** baseline, drop ≤ ⅓
- Only **200 representative rows** reach the model

**Decades-old statistics, doing the heavy lifting!**

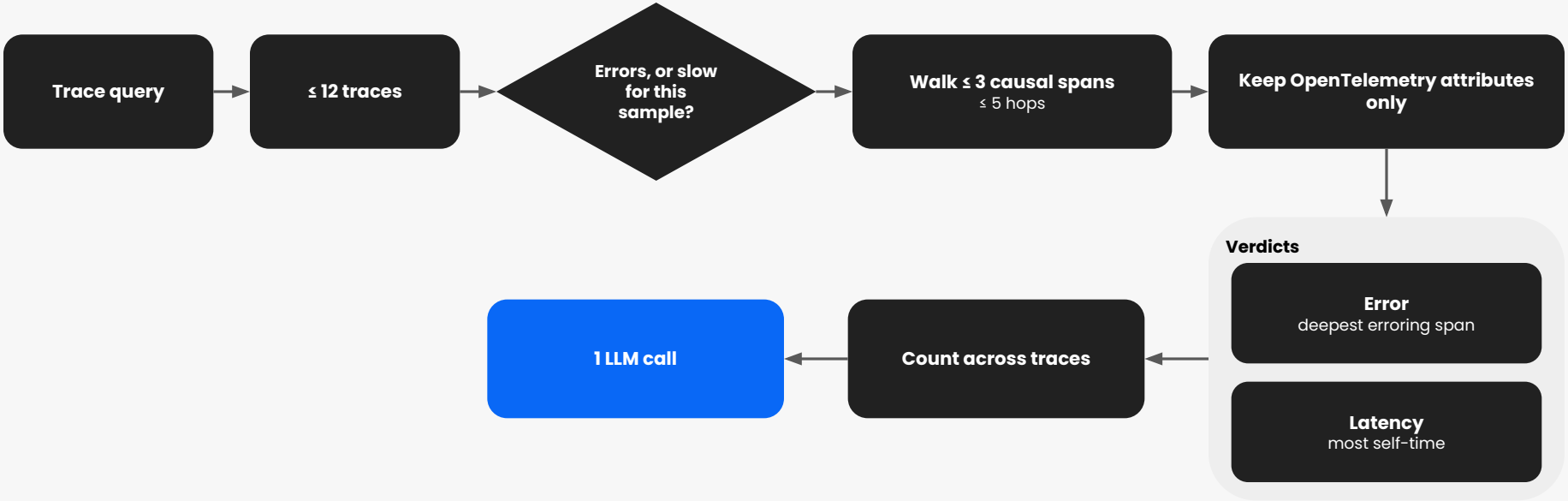





© 2026 eBay. All rights reserved. Confidential and proprietary.

# Traces: Verdicts, Not Guesses

Go LLM




**“Slow” is relative to the sample, not a fixed number!**

© 2026 eBay. All rights reserved. Confidential and proprietary.

# But Can We Trust It?

© 2026 eBay. All rights reserved. Confidential and proprietary.

# The LLM Cites, It Doesn't Author

- Code builds a menu of evidence: **ct-1**, **tt-1**, …
- Behind each handle: the real trace / log ID and the real query
- The model can only order from the menu
- Unknown handle → evidence **dropped**
- Finding and hypothesis IDs are assigned by the server

**It physically can't hand on-call a made-up trace ID!**





© 2026 eBay. All rights reserved. Confidential and proprietary.

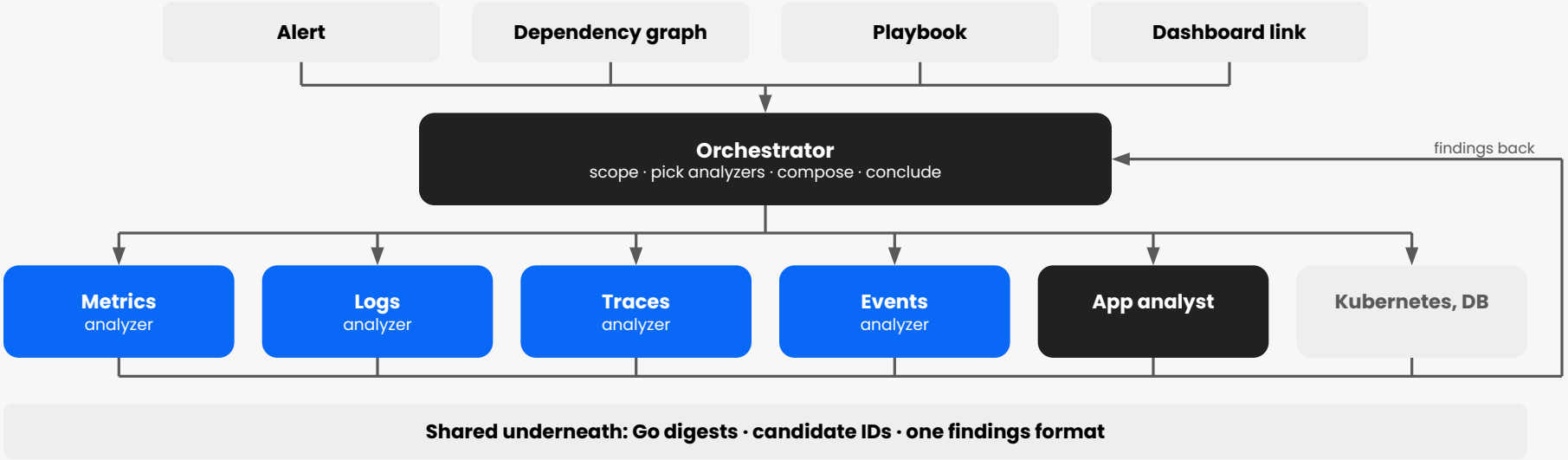
# 04

# What's Next...

© 2026 eBay. All rights reserved. Confidential and proprietary.

# Reusing the Blocks, Not the Dashboard

Dashboard Analyzer




**The dashboard link is one input. The parts underneath are what we reuse.**

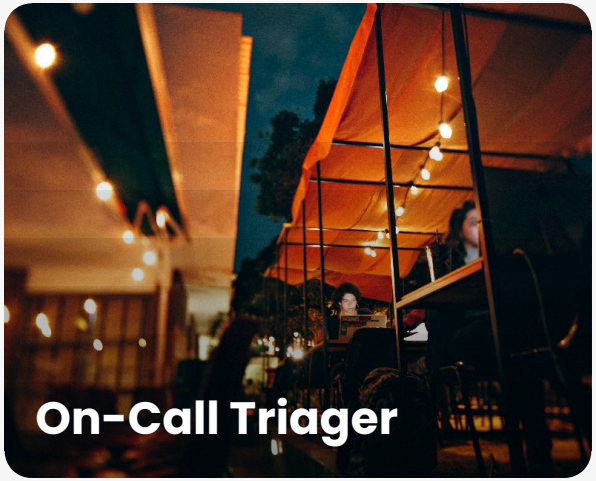
Analyzers and digests run today. The orchestrator and app analyst are in design.

© 2026 eBay. All rights reserved. Confidential and proprietary.

# Toward AI SRE Personas




Starts from: a deploy
Different from before? (baseline)
Where? (metrics + traces analyzers)




Starts from: an alert
Scope: alert, dependencies, playbook
Analyzers chosen by the playbook




Starts from: the findings
Evidence, links, what nobody checked
→ a timeline to start from

**Each persona is a trigger and a playbook on the same blocks, not a new pipeline.**

Direction, not shipped.

© 2026 eBay. All rights reserved. Confidential and proprietary.

# So, What Should You Take Home?

© 2026 eBay. All rights reserved. Confidential and proprietary.

# Three Takeaways

- **Reduce in code; let the LLM narrate**
- **The LLM never authors identity**
- **Budgets, not freedom**, and set your MCP tool parameters explicitly

**Keep your MCP servers. Change what sits in between.**



© 2026 eBay. All rights reserved. Confidential and proprietary.



# Thank You!