

2AM PAGER CHECKLIST

Sheet v1.0 | Levels: runbook v0.1.1 | 2026-10-06

AI pipelines: ten symptoms, the alert level for each, and the first move.

THE FIRST FIVE MINUTES

- 1 Ask what changed at the hour this started: a deploy, a prompt edit, a model swap, a key or quota change.
- 2 If several alerts fire together, find the shared cause first (panel below) before working them one by one.
- 3 Confirm the number on your own logs. Silence is not health: a pipeline that stopped logging looks like a quiet one.

Baseline = trailing 7-day average. Levels are runbook v0.1.1 defaults, not benchmarks: tune them to your own.

THE TEN SYMPTOMS

☐ Cost spike

WARN >2x **PAGE** >3x **spend/1k calls**

One step's model or prompt changed.

First move: Find the spending step. Revert its model or prompt.

☐ Latency p99

WARN >20,000 ms **PAGE** >30,000 ms

The provider is not answering, or a bigger model.

First move: Check the provider status page. Do not just raise the timeout.

☐ Judge failure rate

WARN >5% **PAGE** >10% **fail the judge**

Quality dropped on one step, or the judge moved.

First move: Revert that step. If every step fails, suspect the judge.

☐ Empty output rate

WARN >2% **PAGE** >5% **empty or cut off**

Output cap too low, a filter, or a stop sequence.

First move: Cut off: raise the cap. Empty: read the raw response. Do not blind-retry.

☐ Refusal rate

WARN >1% **PAGE** >2%

A new prompt or model tripped a refusal.

First move: Read the refusals and fix the template. Never retry unchanged.

☐ Schema drift

WARN <98% **PAGE** <95% **valid**

Usually upstream: empty, cut-off or refused outputs.

First move: Fix those first. Quarantine invalid rows, do not load them.

☐ Retry storm

WARN >1.5x **PAGE** >2x **calls/trace**

Retries are amplifying the throttling.

First move: Backoff with jitter. Cap 3 to 4 attempts. Add a circuit breaker.

☐ Cache hit collapse

WARN <0.5x **baseline** **no page**

The prompt prefix changes between calls.

First move: Stable text first, per-request text last.

☐ Provider errors

WARN >0.5% **PAGE** >1% **429 or 5xx**

You are throttled, or the provider is down.

First move: Check the status page. Lower concurrency, back off, fail over on 5xx.

☐ Token per call

WARN >2x **no page**

Prompt bloat, the earliest and cheapest sign.

First move: Diff the template. Cut pasted context and repeated instructions.

WHEN SEVERAL FIRE AT ONCE

cost spike + token per call + cache collapse One prompt or model change. Start at cost spike.

provider errors + retry storm + latency p99 Throttling made worse by retries. Start at provider errors.

judge + refusals + empty output + schema drift One step's output quality dropped. Start at judge failure rate.

WHAT THE FULL KIT'S CHECK PRINTS

```
{"check": "cost_spike", "status": "page", "value": 5.301, "threshold": 3.0, ...}
```

From the runbook's sample incident day: spend per 1,000 calls was \$9.29 against \$1.75 for a normal week. Try the idea on your own export with `check_cost_spike_demo.py` (included, free).

This page is the triage table. The full 2AM Pager Runbook for AI Pipelines (\$29) is the diagnosis: ten runnable checks, three narrowing queries per symptom, the full fix list, escalation rules, and a sample database with an incident day.

Not covered here: which step, model or prompt in your system is at fault. Assumes you track, or can query, these ten numbers.

dataengineerai.gumroad.com/l/godkks

Questions: dataengineering.with.ai@gmail.com