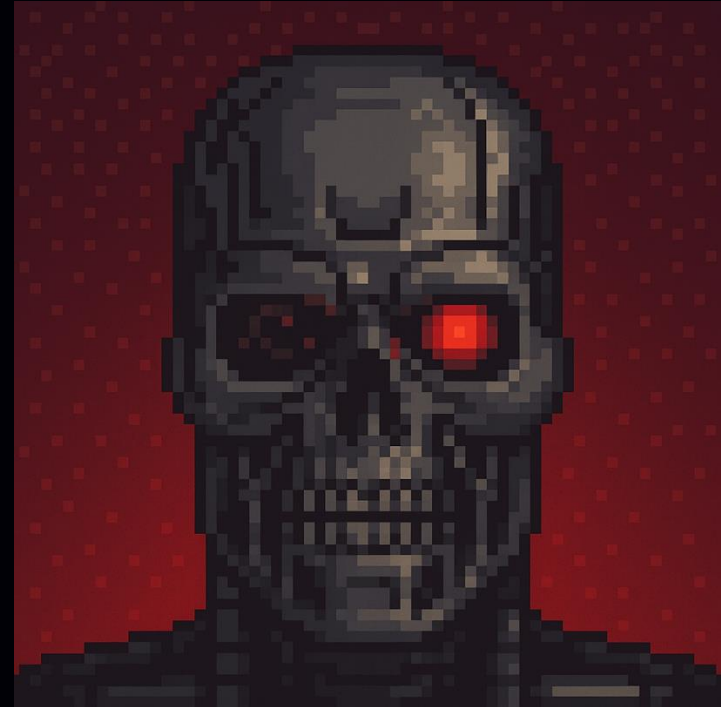# I'll Be Backoff: Benchmarking AI-Powered Platform Engineering Tools

Annie Talvasto,

Sr. Manager, Product Marketing & Community

Upbound

# AI & Platform Engineering

# Why should you care?

# What are the options?


WORK MODE

# Landscape: 4 Major Categories

**1. AI SRE (Incident Intelligence)**

Tools that reason, triage, investigate, and suggest or coordinate fixes.

**2. AI Platform Engineering / SDLC Automation**

Tools that automate workflows, tickets, build-test-release, scorecards, and developer experience.

**3. Infrastructure Automation (IaC & Control Planes)**

AI for infra mutation and enforcement.

**4. Agent Runtimes (Agentic Infra)**

Where agents actually run, the execution substrate for agentic workloads.

# AI SRE
# (Incident
# Intelligence)

# Datadog Bits AI SRE

- Fully autonomous investigation inside Datadog

- Telemetry-native: logs, metrics, traces, Watchdog

- Incident summaries, coordination, postmortems

- Best for orgs already on Datadog

# Phoebe

- High-accuracy log/trace/query reasoning

- Predicts incidents and generates PR fixes

- Builds knowledge graphs for faster debugging

- Vendor-agnostic read-only ingestion

# Cleric

- Transparent hypothesis-driven investigations

- Shows reasoning chains and validated evidence

- Engineer collaboration for fixes

- Read-only by design

# Resolve AI

- Autonomous incident investigation + RCA

- Cross-system evidence collection

- Fix proposals + human approval loop

- Multi-agent orchestration

# Traversal

- Evidence-backed RCA, alert ranking, fast diagnosis

- Designed for on-prem / high-security teams

- Fix suggestions + postmortem automation

- Not focused on SDLC automation

# AI Platform Engineering / SDLC Automation
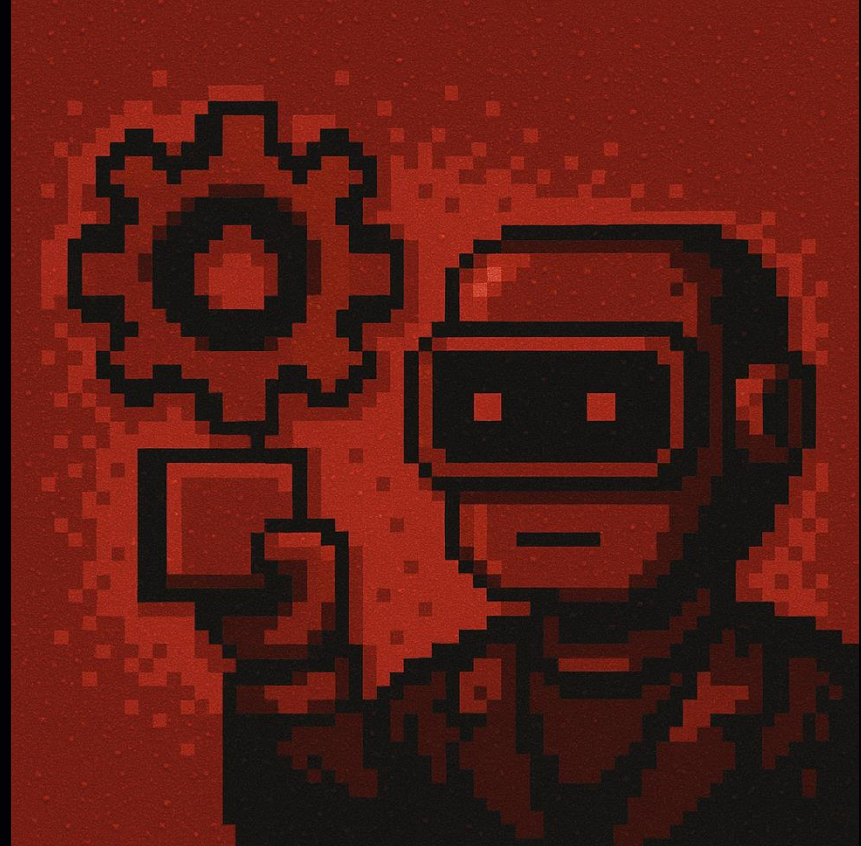
# Port

- Developer portal with context lake + scorecards

- AI agents automate tickets, actions, and workflows

- Strong SDLC catalog visibility

- Not an incident/RCA tool

# SRE AI

- Broad AI across testing, CI/CD, release, security

- Ticket/PR automation, risk detection

- Great for hybrid/remote engineering orgs

- Moderate depth in RCA

# Harness AI SRE

- Incident triage + RCA via change intelligence

- Runbook automation across the Harness platform

- Strong SDLC-wide visibility

- Best value if using Harness CI/CD

# Infrastructure Automation (IaC & Control Planes)
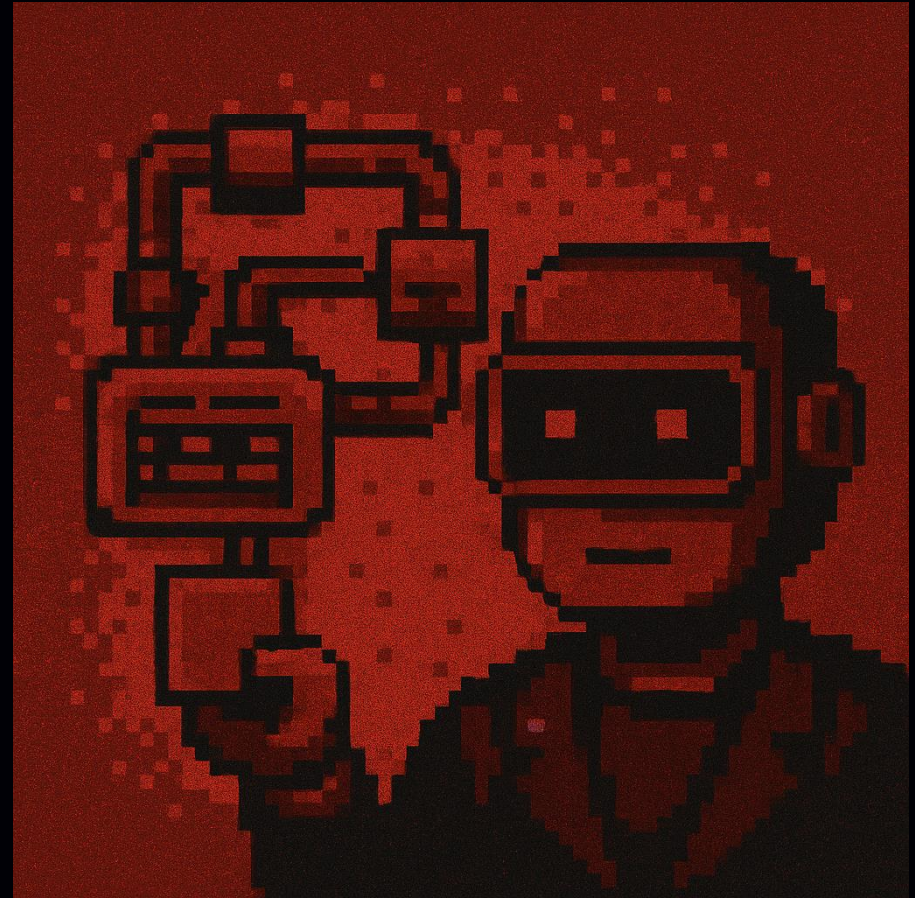
# Pulumi Neo

- AI agent for infrastructure-as-code changes

- Generates PRs, upgrades, fixes, compliance updates

- Strong governance via previews + policies

- Deep Pulumi ecosystem integration

# Upbound

- Declarative, self-healing infra via Crossplane 2.0

- Deterministic APIs for humans + AI agents

- Policy enforcement + drift correction

- Foundation layer; not an SRE/RCA system

- Build on top of the trusted reconciliation technology/controllers as Kubernetes

# Azure SRE Agent

- Azure-native autonomous remediation agent

- Always-on monitoring + active incident fixes

- Works only on Azure resources

- Infra-level autonomy, limited RCA reasoning

# Agent Runtimes
(Agentic Infra)

# Kagent

- MCP/ADK agents running inside Kubernetes

- Cluster-native workflows, tools, tracing

- Extensible via CRDs + open standards

- Not an SRE/RCA product, it's the runtime

| Tool | Primary Domain | What It Actually Does | Best For |
|---|---|---|---|
| **Datadog Bits AI SRE** | Native observability-driven SRE | Fully autonomous investigations, incident coordination, summaries, postmortems | Organizations already on Datadog |
| **Phoebe** | Proactive + reactive AI troubleshooting | Predict issues, RCA, PR fixes, high-accuracy log/search models | High-scale orgs with rich observability pipelines |
| **Cleric** | Explainable AI SRE | Hypothesis-driven RCA, structured reasoning, human-collab debugging | Teams needing transparent, explainable agents |
| **Resolve AI** | AI SRE (multi-agent) | Incident investigation, RCA, code/infra context, fix proposals | SRE teams needing deep incident reasoning |
| **Traversal** | AI SRE with privacy focus | Alert ranking, RCA, remediation suggestions, postmortems, on-prem/bare-metal friendly | Enterprise SRE teams with compliance constraints |
| **Port** | AI-enhanced IDP + workflow automation | SDLC catalog, scorecards, ticket automation, custom AI agents | Large engineering orgs with complex workflows |
| **SRE.ai** | Broad engineering automation | Testing, CI/CD insights, SDLC orchestration, predictive workflows | Hybrid/remote teams needing lifecycle automation |
| **Harness AI SRE** | Incident mgmt + SDLC change correlation | Triaging, RCA via change intelligence, runbook automation | Companies using Harness CI/CD |
| **Pulumi Neo** | AI for IaC automation | Infra code generation, upgrades, compliance fixes, PR automation | Platform engineers using Pulumi |
| **Upbound** | Intelligent control plane (foundation) | Declarative infra APIs, drift correction, policy enforcement | Org building AI-native infra & CPaaS |
| **Azure SRE Agent** | Cloud-native auto-mitigation | Azure resource monitoring + automated fixes via AAUs | Teams fully on Azure needing auto-reliability |
| **kagent** | Kubernetes-native agent runtime | Run MCP/ADK agents *inside* clusters; agent workflows; observability | Platform teams wanting open standards for agentic infra |

| Tool | Grounding Source | Strength |
| --- | --- | --- |
| **Bits AI (Datadog)** | Full Datadog telemetry graph | Very strong telemetry-native grounding |
| **Phoebe** | Proprietary cross-silo log/trace/search models | High-accuracy reasoning across systems |
| **Cleric** | Logs, metrics, code changes | Explainable evidence-based grounding |
| **Resolve AI** | Logs, metrics, traces, code, infra maps | Broad multi-source grounding |
| **Traversal** | Telemetry + deploy history | Precise incident signal grounding |
| **Port** | SDLC metadata, org catalog, workflows | Strong organizational context |
| **SRE.ai** | CI, code, tests, SDLC process data | Broad SDLC context |
| **Harness AI SRE** | CI/CD + change context | Good for SDLC-to-incident correlation |
| **Pulumi Neo** | IaC repos, Pulumi state, policies | Strong code-level grounding |
| **Upbound** | Control plane state, CRDs, policies | Deterministic infra grounding |
| **Azure SRE Agent** | Azure Monitor, resource health | Strong Azure-native grounding |
| **kagent** | K8s resources, MCP tools, ADK | Strong cluster-native grounding |

| Tool | Autonomy Level | What It Can Autonomously Do |
| --- | --- | --- |
| **Bits (Datadog)** | High for RCA, low for changes | Autonomous reasoning + incident coordination |
| **Phoebe** | High | Predictive fixes, PR generation |
| **Cleric** | Medium | RCA + explanations, not infra writes |
| **Resolve AI** | Medium-high | Multi-agent RCA + fix proposals |
| **Traversal** | Medium-high | RCA + remediation suggestions |
| **Port** | Medium | Execute actions, workflows, route tickets |
| **SRE.ai** | Medium | Testing, release gating, risk detection |
| **Harness AI SRE** | Medium | Execute runbooks + correlate changes |
| **Pulumi Neo** | Medium | Generate PRs, run previews, propose changes |
| **Upbound** | High (declarative) | Continuous reconciliation + drift correction |
| **Azure SRE Agent** | High (infra only) | Auto-mitigations in Azure |
| **kagent** | Varies by CRD | Agents execute workflows *in-cluster* |

| Tool | Strength Highlights | Limitations |
| --- | --- | --- |
| **Bits AI** | Deepest telemetry-native grounding | Datadog-only ecosystem |
| **Phoebe** | Strong proactive detection + accurate PR fixes | Not an incident coordinator |
| **Cleric** | Explainable reasoning chain | No automated infra writes |
| **Resolve AI** | Multi-agent RCA with code/infra intelligence | No infra drift control; not deterministic |
| **Traversal** | Fast RCA + enterprise privacy posture | Not multi-cloud infra automation |
| **Port** | Strong workflow engine + context lake | Not RCA/incident-focused |
| **SRE.ai** | Full SDLC-level automation + testing intelligence | Wide scope but shallow RCA depth |
| **Harness AI SRE** | Change intelligence inside CI/CD | Requires Harness stack |
| **Pulumi Neo** | Tight IaC grounding, safe PR automation | Not an SRE system; no observability reasoning |
| **Upbound** | Declarative control planes; infra correctness | No RCA, no incident workflows |
| **Azure SRE Agent** | Auto-remediation for Azure workloads | Azure-only; limited RCA depth |
| **kagent** | Open-standards agent runtime for K8s | No RCA or observability; needs K8s maturity |

# Landscape: 4 Major Categories

**1. AI SRE (Incident Intelligence)**

Tools that reason, triage, investigate, and suggest or coordinate fixes.
Examples: Datadog Bits, Phoebe, Cleric, Resolve AI, Traversal

**2. AI Platform Engineering / SDLC Automation**

Tools that automate workflows, tickets, build-test-release, scorecards, and developer experience.
Examples: Port, SRE.ai, Harness AI SRE.

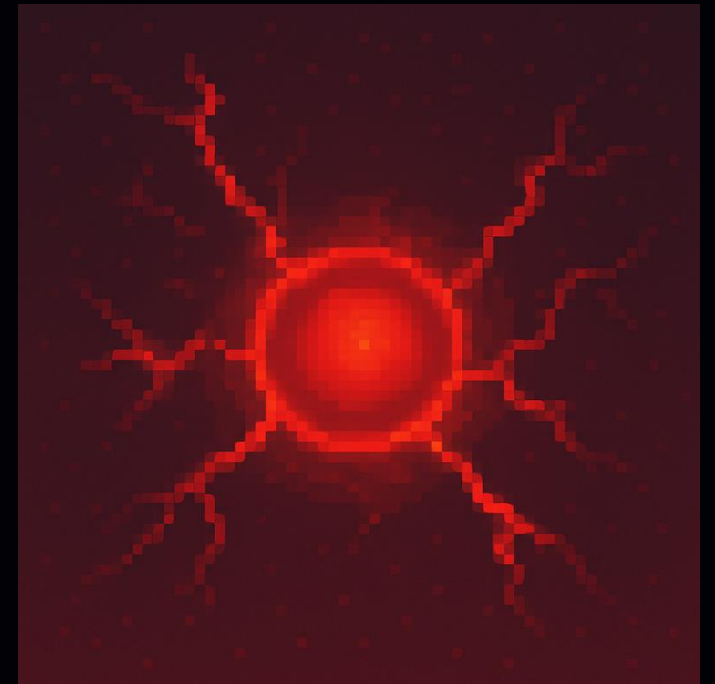**3. Infrastructure Automation (IaC & Control Planes)**

AI for infra mutation and enforcement.
Examples: Pulumi Neo, Upbound, Azure SRE Agent.

**4. Agent Runtimes (Agentic Infra)**

Where agents actually run, the execution substrate for agentic workloads.
Example: kagent

# Resources

- Pick what makes most sense for you!

- Intelligent Control Plane
  - https://www.upbound.io/intelligent-control-plane

- SREday sessions
  - **AI & SRE: Embracing the Ultimate Frontier of Automation**
  https://youtu.be/7VXZBLROglM?si=2ezbKagZTsR8b3C5
  - **The Rise of SRE AI Agents**
  https://youtu.be/xWNl4WLxXcw?si=tJkJsOvxEoHsAckY
  - And many more!

# Thank you!