chronosphere + arize

SOLUTION BRIEF

Chronosphere & Arize AI

Combine LLM monitoring with full stack observability to keep AI services accurate and reliable at scale.

THE CHALLENGE

User-impacted AI issues slip through without connected visibility

Modern engineering teams need to detect model drift, degraded accuracy, or bias issues before customers feel the pain. Model monitoring provides early warnings, but without backend context, teams risk chasing symptoms instead of solving root causes. AI-driven issues are often discovered too late. Model performance signals alone can't pinpoint backend or infrastructure dependencies. Teams troubleshoot across siloed ML and DevOps tools, slowing incident resolution.

THE SOLUTION

Full coverage from model performance to Kubernetes cluster

Arize continuously validates AI agent health, model outcomes, data quality, and fairness, appropriateness and accuracy across development and production environments. Chronosphere ingests those insights as OTel traces, metrics, and events to provide backend context. Engineers detect issues early, correlate AI outcomes, and ML drift with infrastructure signals, and remediate faster.

KEY BENEFITS

Catch inference issues early. Identify data drift, degraded accuracy, or fairness regressions before customers are impacted.

Correlate model behavior with backend systems. Connect AI performance signals to Kubernetes workloads, APIs, or services.

Control your AI data. Cut through high volumes of observability data with targeted model insights from Arize.

How it works

The integration provides a continuous loop from ML detection to backend troubleshooting:

Use OTel SDKs to monitor real-world AI signals with Arize — model performance, drift, fairness, and data quality.

Trigger alerts instantly when predictions degrade, features drift, or outliers appear.

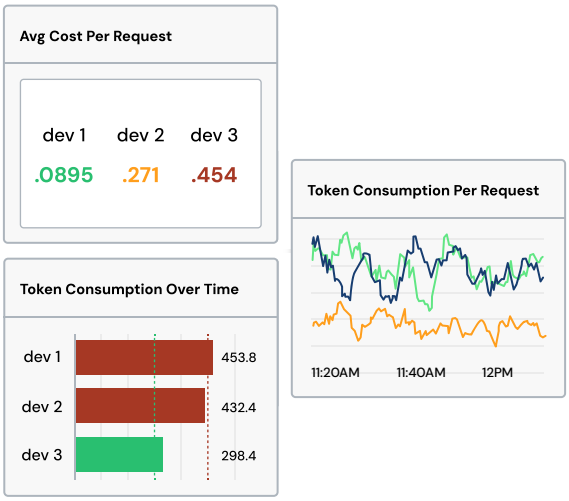
Forward AI monitoring telemetry into Chronosphere, unifying them with backend observability data.

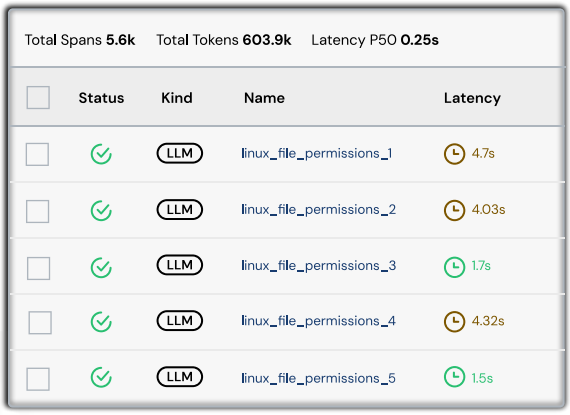
Correlate issues across the stack – from model behavior to Kubernetes pod metrics and logs.

Remediate faster with Chronosphere's high-context insights and Arize's debugging artifacts.

See AI performance and cost in context

Chronosphere unites AI-specific telemetry from Arize with traditional observability data — RED metrics, infrastructure monitoring, and MELT telemetry — in one view. SREs gain real-time visibility into LLM token consumption, GPU usage, hallucinations, drift, and bias alongside service health metrics. With full-stack correlation, teams can detect spikes in hallucinations or rising GPU costs and resolve root causes faster, ensuring every AI service runs efficiently and reliably at scale.





Trace every token, reveal every dependency

By instrumenting traces with the OpenInference SDK from Arize AI, Chronosphere exposes every LLM request — inputs, outputs, and evaluation results — within the broader service map. Engineers can visualize where hallucinations or drift emerge in the request path and how they affect dependent systems. With trace-level metrics and trend tracking, teams gain granular understanding of model performance, latency, and reliability, empowering precise root cause analysis for complex AI-powered environments.

Take control of your observability.

Connect with a Chronosphere expert to see how you can optimize costs, reduce noise, and resolve issues faster.

Book a demo

About Chronosphere

The observability platform built for control in the modern, containerized world. Recognized as a leader by major analyst firms, Chronosphere empowers customers to focus on the data and insights that matter to reduce data complexity, optimize costs, and remediate issues faster.

About Arize

Arize AI is an AI and agent engineering platform. Arize's tools allow teams to quickly detect issues when they emerge, troubleshoot why they happened, and improve overall performance across both traditional ML and AI agent engineering.