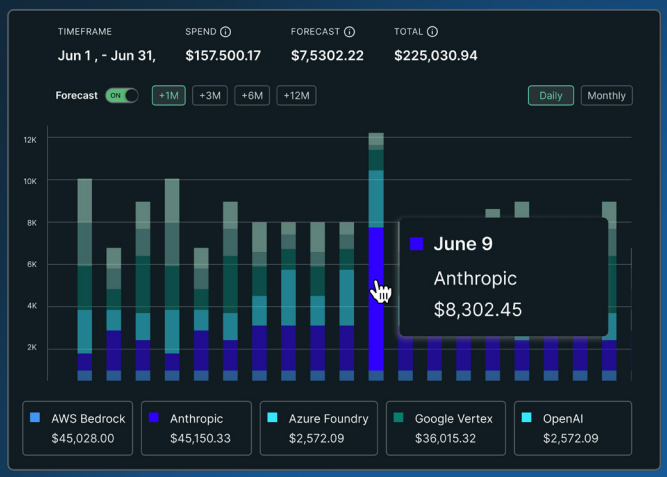
FinOps+ by Kion

FinOps for AI

Get started with AI tokenomics: Visibility, cost control, and governance through automated enforcement of AI token usage and spending.



CHALLENGE

As AI adoption expands, so does its cost footprint

As organizations expand their use of AI, it becomes increasingly difficult to gain a clear view of total AI spend both across cloud providers (including AWS Bedrock, Azure AI Foundry, Google Enterprise Agent Platform, and OCI) as well as native AI providers such as OpenAI and Anthropic.

Elevate your FinOps practice

Kion brings FinOps discipline to AI spend, providing **visibility into model and token costs, enabling accurate cost allocation, and enforcing budgets and guardrails** to keep spending under control.

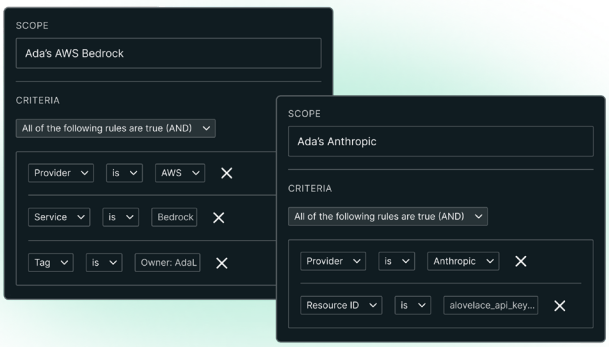
AI COST MANAGEMENT WITH KION

- Unified Visibility
- Proactive Cost Optimization
- Automated Governance Across Cloud, AI, & Beyond

SOLUTION

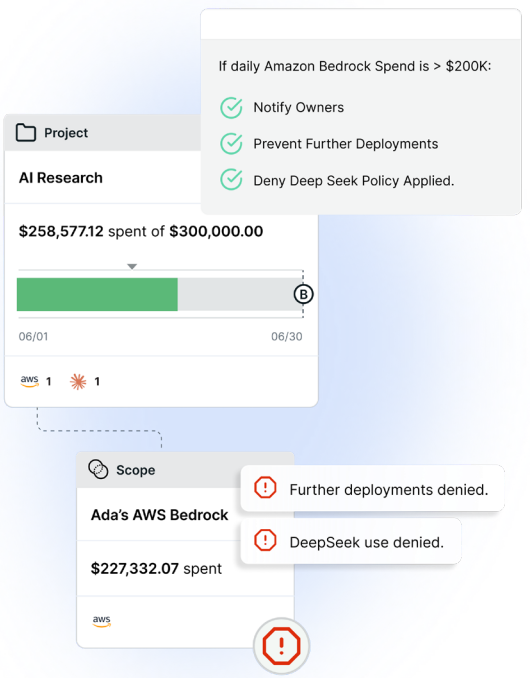
Accurately Attribute AI Costs

With Kion's Scopes feature, you can flexibly group and allocate AI costs to the teams, projects, and departments where it is being used, providing insight into model utilization, token costs, and the underlying token economics behind AI spend.



Visibility Alone Won't Prevent Overspending

Understanding where AI spend is occurring is only the first step. Without budgets, guardrails, and enforcement, costs can continue to grow unchecked.



Set Budgets & Preventative Financial Guardrails

With budgets and guardrails, you can notify stakeholders as AI token costs and usage approach defined thresholds, or automatically enforce controls to stop additional activity, preventing cost overruns before they occur.

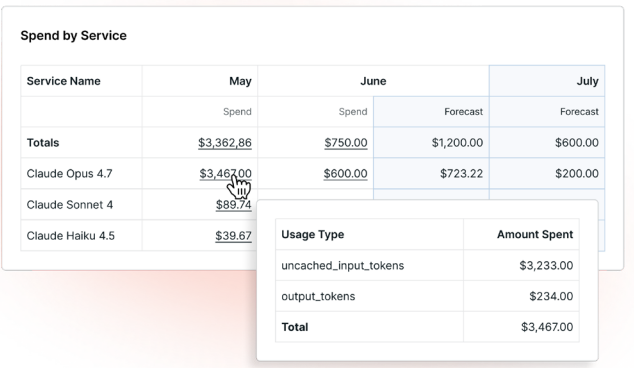
AI Compliance Guardrails

In addition to financial guardrails, Kion provides compliance guardrails that help organizations enforce AI governance policies and reduce risk.

Prevent access to disallowed models such as DeepSeek, Fable 5, or others, to ensure AI usage remains aligned with your organization's security, compliance, and governance standards.

Granularly Track, Analyze, & Forecast AI Spend

Analyze token cost by model, provider, input vs. output, and cached vs. uncached to identify optimization opportunities. Track usage patterns across tokens and GPUs. Manage GPU capacity commitments with the same rigor as reserved cloud capacity.



Learn more or schedule a demo online today.
www.kion.io

Kion