

Case Study: Governing the AI Spend Surge at a Consulting Firm

A global advisory firm of 3,000 consultants delivers strategy and implementation work across hundreds of concurrent client engagements. Like most professional services firms, it sells hours, tracks them meticulously, and prices a large share of its work on a fixed-fee basis.

3,000
CONSULTANTS

\$45,000+
MONTHLY AI SPEND

Hundreds
CLIENT PROJECTS

30 days
TIME TO VALUE

Challenge

AI code-assist and cowork tools spread quickly across the firm, with spend climbing from zero to \$45,000 a month in a matter of months. Nobody could see which consultant used which tool, for how much, or on whose behalf, so none of the costs could be attributed to a project or a client. The firm measured every hour and priced much of its work on a fixed basis, but AI had become a substantial input in the business model and it could not measure or account for it.

Solution

- Collected AI usage from every source: cloud AI, a direct connection to the AI provider's API, and an OTEL collector for code-assist and cowork tools.
- Built a virtual-tag taxonomy mapping each consultant to their practice, team, project and client, which resolved historical and new usage.
- Gave each manager a view of AI usage and spend across their own teams and projects.
- Rolled usage up into chargeback by client and project.

Benefits

- **AI Return on Investment:** AI spend is compared against the consulting hours it replaces.
- **Cost answers in seconds:** the firm queries 1,850 metrics and dimensions in plain language.
- **Chargeback:** the firm recovers AI spend from the clients and projects it was used for.
- **Productivity Insights:** managers see usage per consultant and spot power users and coaching needs.
- **Scalable Governance:** new projects, teams and tools inherit the taxonomy automatically, with no rework as the firm grows.

Cloud & AI Cost Management

Cloud and AI spend is now one of the fastest-growing lines on the technology cost base, and one of the least understood. Costs sit in shared environments that cannot be tagged, arrive as invoices that say what you paid but not why, and rarely resolve to the team, project or customer that drove them. Agents make it harder still, running continuously and decoupling spend from headcount.

CostPerform Cloud & AI ingests cost and usage data from every major cloud and AI provider, allocates it with audit-grade logic to the teams, clients and products behind it, and turns that visibility into showback, chargeback and unit economics. Trusted by over 200 organizations, including British Telecom, ING bank, and the USPTO, CostPerform delivers unparalleled financial clarity and cost optimization.

costperform.com/demo



w. costperform.com
t. +31 (0) 85 023 2114
e. info@costperform.com

CostPerform Cloud & AI for Enterprise

- **True Unit Economics:** Break spend down to the token, model, agent or workload, so you can measure the return on AI and cloud investment.
- **Allocate and Charge Back Instantly:** Split shared GPU pools, model inference and agent activity by team, business unit or client, so shared spend lands where it was actually used.
- **Proactive Governance:** Set budget thresholds by team, project or model, and catch the deviations you didn't predict with anomaly detection.
- **Connect Every Platform:** Native connections to 15+ cloud and infrastructure platforms, including AWS, Azure, Google Cloud, Oracle, Kubernetes, so nothing sits outside the model.
- **Capture Every AI Signal:** Ingest usage from all standard enterprise AI tools, including Anthropic, OpenAI, Gemini, Bedrock, and Vertex, so no spend goes unattributed.
- **Prove Every Number:** Trace every allocation back to transparent, versioned logic that holds up with customers and auditors.
- **See the Whole Cost Picture:** Carry the same model from Cloud and AI through to bill of IT, shared services and enterprise-wide profitability.

costperform.com/demo

CostPerform USA

+1 410 385 5325

infousa@costperform.com

CostPerform The Netherlands

+31 (0) 85 023 2114

info@costperform.com

