



Enterprise AI adoption has outpaced enterprise AI governance. **API keys sprawl across teams. Spend is invisible. Policy is unenforceable.** Traditional API gateways were built for HTTP — not tokens, prompts, models, or agents. The result: security exposure, runaway costs, and zero auditability across every AI workload in production.

SOLUTION OVERVIEW

MajorDomo AI Gateway is an enterprise AI gateway and governance platform that sits in the path of AI interactions to secure access, enforce policy and deliver full observability across every model, agent and MCP interaction. Its federated Director/Worker architecture enables centralized governance, security, and operational control while allowing teams to use the models, agents, and deployment environments best suited to their needs. Model-agnostic and multi-cloud ready, MajorDomo provides consistent security, compliance, visibility, and policy enforcement across the enterprise AI ecosystem.



CORE CAPABILITIES

Access Control & Key Management - Workspace and project-scoped virtual keys with full lifecycle management - Agent-specific key isolation — no shared credentials across workloads - RBAC with SSO/IdP integration and provider key masking - Centralized revocation and rotation without disrupting consumers	Policy & Governance Engine - Model allow/deny lists enforced by team, project, or workload - Token limits and budget caps with hard per-workspace enforcement - Policy-driven model routing and fallback models - Prompt injection, jailbreak protection and PII detection
Observability & Audit - Session-level traces across tool calls, reasoning steps, and MCP interactions - Immutable audit logs for compliance and regulatory reporting - Real-time cost telemetry by model, team, project, and agent - Configurable alerting on spend, latency, and error rate anomalies	Cost Governance & Chargeback - Per-workspace spend dashboards with real-time visibility across all providers - Automated budget enforcement — hard stops before any cost overage - Chargeback reports exportable by team, cost center, or business unit - Token analytics with model-level attribution and spend forecasting

MAJORDOMO VALUE PROPOSITION

Product & Engineering Teams - Integrate AI capabilities without exposing raw credentials or API keys to consumers - Intelligent model routing across 100+ endpoints — route by cost, latency, or capability without refactoring integrations - Built-in rate limiting, retries, and per-project cost guardrails out of the box	Platform & Security Teams - Complete audit trail of every model, agent, and MCP interaction across the enterprise - Automatic PII detection and masking on every request and response before data leaves the perimeter - Enforce access policies and guardrails centrally — air-gap and private cloud ready for regulated environments	Leadership & Business - Eliminate AI spend surprises with token-level attribution and forecasting across all teams - Accelerate enterprise AI adoption with governance guardrails that don't slow delivery - Meet compliance requirements across every AI workload from a single control plane
---	---	---

100% Observability across all AI traffic	Zero Raw credentials exposed to consumers	100+ Model endpoints supported	Air-Gap Data-sovereign deployment ready	MCP Native MCP server governance & security
--	---	--	---	---