

Cequence AI Gateway

Security, governance, and control for agentic AI

Cequence built the industry's most effective bot management solution and now secures over 10 billion API calls and 200 million agentic interactions every day. The behavioral intent engine at the core of our platform separates good users, good bots, and fraudulent actors by analyzing intent, not just identity. The AI Gateway extends that proven technology to AI agents: every agent gets a defined job, every action gets verified against that job continuously, and trust is never assumed from only a credential.

Because enforcement lives at the AI Gateway rather than in the model or the endpoint, governance is centralized no matter where agents run: managed devices, cloud platforms, or SaaS agent services like ChatGPT workspaces and Agentforce. Frontier, open-weight, and self-hosted models are all protected. Identity gets the agent in; the AI Gateway governs what it does next and verifies every action against its role.

What Makes the Cequence AI Gateway Different



Agentic Zero Trust Architecture

AI Gateway authenticates agents and then authorizes every action they take. The gateway enforces policy inline with the request path, for the entire session, on every tool call. Dr. Chase Cunningham and Anthropic independently described this as the right approach; it's the same architecture Cequence had already built.



Least Privilege Access with Agent Personas

A plain-English job description ensures agents only get the specific tools, API calls, LLMs, skills, and data they actually need — everything else stays off-limits. Personas work immediately from the first tool call, with no learning period required.



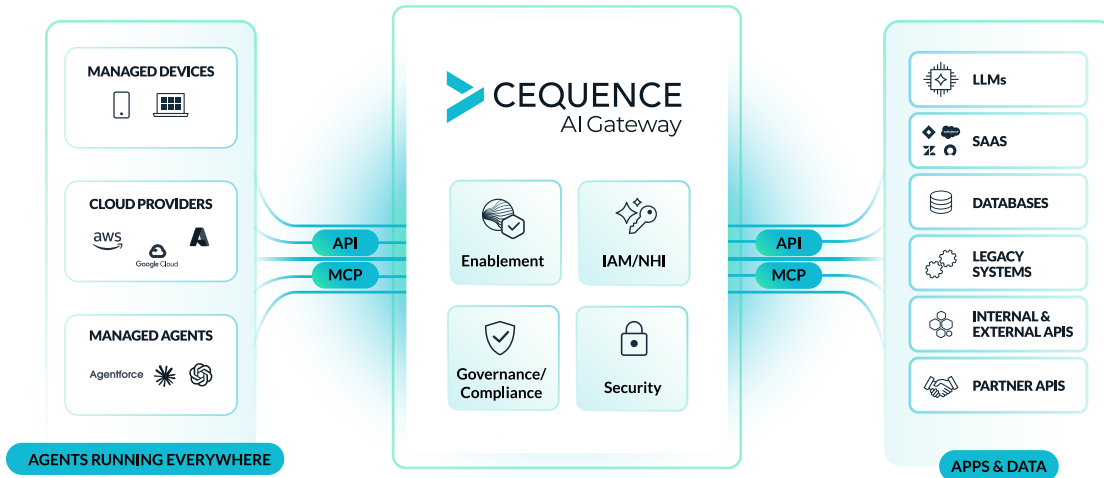
Behavioral Intent Monitoring

The Cequence behavioral intent engine analyzes the sequence of actions an agent takes, not simply one call at a time, against per-persona baselines and halts actions that deviate from that baseline.



Sensitive Data Protection

Inline detection, masking, redaction, and blocking on every request and response, with more than 100 built-in detection types supporting PCI-DSS, PHI, SOC 2, and HIPAA compliance.

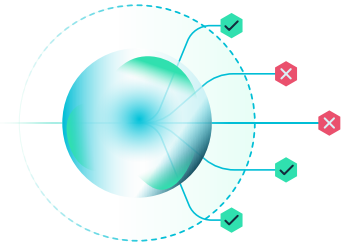


The Cequence AI Gateway provides the security and governance enterprises require to confidently deploy agentic AI workflows at scale.

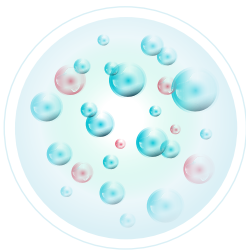
AI Gateway Capabilities

Agentic AI Governance

Agent Personas. Every agent gets a job description that dictates tool access, governed by a central Skill Registry that defines which skills bind to which personas. Each persona maps to a single virtual endpoint and enforcement happens inline at the AI Gateway. Agent Personas expose only the tools, APIs, LLMs, skills, and permissions the job requires, so the agent receives a short, relevant list instead of hundreds in every request, which cuts token consumption and improves performance as the model selects the right tool the first time.



Sensitive Data Protection. Actions often pass permission checks and still exfiltrate sensitive data, so AI Gateway inspects every AI agent request and response at runtime. Monitor, redact, and block sensitive data across more than 100 out-of-the-box detection types. Easily integrates with existing DLP infrastructure.

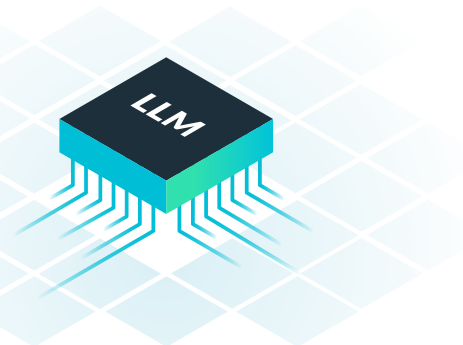


AI Discovery. You can only govern what you're aware of. Cequence provides visibility into existing agent and AI use, discovering sanctioned and shadow AI alike. It surfaces every agent, MCP server, and LLM provider enterprise-wide from existing SIEM logs.

Monitoring and Visibility. Every request records agent, human, tool, time, and outcome, producing the comprehensive audit trail that compliance frameworks and incident response demand. Data and logs export in OTEL format for SIEM and GRC integration.

Enterprise MCP and API Registries. The MCP Registry provides a trusted catalog of vetted MCP servers built from official application APIs as well as custom MCPs for your own applications, all centrally provisioned with IAM capabilities. The AI Gateway handles MCP protocol revisions, so no code changes are required as the standard evolves. The API Registry lets agents call approved business systems without holding origin credentials.

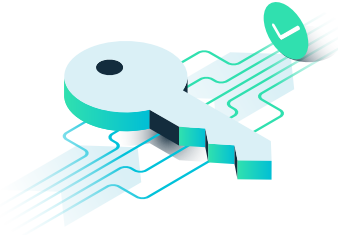
LLM Registry and Prompt Guard. The LLM Registry provides a governed connection to each approved LLM provider, manages credentials and model access, and applies token rate and spending limits. Prompt Guard protects against prompt injection including jailbreak and system-prompt-extraction patterns before they reach the model.



Security

Behavior Anomaly Detection. Behavioral baselines per persona, per user, and per tool catch the patterns single-call inspection misses. An agent reading 200 tickets in a row before opening an email tool would pass every individual permission check; the sequence here is the violation, and the gateway throttles or halts it mid-action based on policy.

Guardrails and Rate Limits. Per-agent, per-tool rate limits with circuit breakers for runaway loops, automated tool risk scoring, and network controls including IP CIDR restrictions, geo-fencing, and IP pinning which requires tokens to be used from the IP address to which they were issued, per persona.



Credential Isolation. The agent authenticates to the gateway with a credential that grants access to the gateway and nothing else. The agent and the model never see a backend secret, so a compromised agent or manipulated prompt cannot leak credentials to your systems.

Enterprise Authentication. Multi-IdP support with OAuth 2.1, PKCE, dynamic registration, OIDC integration, and support for legacy authentication, enforcing organizational identity policy for both human and non-human identities.

Enablement

No-Code MCP Server Creation. Upload an existing OpenAPI or Swagger spec, or choose from discovered application APIs, and select the endpoints to expose as tools. The gateway generates an MCP server in minutes with no coding, and every server it creates inherits persona scoping, behavioral monitoring, and guardrails. Deploy fully managed in the Cequence Cloud or self-managed with a Helm chart.

Skill Registry. The Skills Registry provides a central catalog to distribute enterprise-approved instructions and operating patterns. It makes it easier for agents to use approved Skills and for the organization to track and audit agentic actions.



Proven in Production

A major global telecom found that a legitimate developer's coding agent ran a task over the weekend, encountered dependencies that prevented it from completing the task, and then attempted 2.5 million tool calls trying to circumvent the roadblock: fabricating file paths, manipulating file SHAs, and probing for write access. Every credential was valid the entire time. The Cequence AI Gateway detected the abnormal behavior, alerted security teams, and generated the complete audit trail and root-cause report for the investigation.

Built for the Enterprise

Deploy as full SaaS with a dedicated tenant per customer, or on premises where your sensitive data never reaches the Cequence control plane. RBAC, continuous environment monitoring, and discrete pre-prod and prod modes are standard, and the platform is ISO 27001 certified, PCI DSS compliant, and maintains a SOC 2 Type II attestation report. Integration with Cequence API Security and Bot Management adds enhanced API specs that boost agent accuracy plus protection from agent-fueled attacks, abuse, and fraud.

Summary

Every enterprise AI roadmap points the same direction: more agents, more autonomy, and access to more applications and data. The Cequence AI Gateway makes that trajectory governable: provision each agent like a new hire with a default-deny role, watch its behavior against that role in real time, and stop the agents whose actions stray from the job description. Cequence co-authored the CIS Model Context Protocol Companion Guide and co-chairs TM Forum's AI-Native Blueprint Initiative, helping define the standards for securing agentic interactions. Whether you are evaluating agent platforms or already running agents in production, Cequence delivers the security, governance, and control that organizations require.