gruve

SOLUTION BRIEF - PROFESSIONAL SERVICES

AI Gateway Design & Deployment for AI-Accelerated Services: Transform enterprise APIs into secure, agent-ready capabilities

01

Business problem

Organizations are under intense pressure to deploy AI agents that interact directly with enterprise systems such as CRM platforms, databases, ticketing systems, and business applications.

However, most enterprise APIs were never designed for autonomous AI consumption. As a result, organizations face major challenges including:

- **Lack of secure, standardized exposure of APIs to AI agents:** Existing REST APIs lack authentication, authorization, and security controls appropriate for autonomous AI consumption
- **Insecure MCP servers created by developers outside governance processes:** Developers download untrusted MCP servers from the internet or create custom servers without security review, creating backdoors into production systems
- **Manual, time-consuming custom development to enable agent workflows:** Each AI use case requires months of custom integration development, delaying time-to-market and consuming engineering resources
- **Inconsistent authentication and authorization across AI tool calls:** Ad-hoc approaches to securing AI-to-application interactions create security gaps and compliance risks
- **Absence of guardrails to prevent sensitive data exposure:** No centralized controls preventing AI agents from accessing or exfiltrating PII, credentials, or business-critical data
- **No centralized visibility into AI tool usage and workflow execution:** Security teams lack telemetry showing what AI agents are doing, which APIs they're calling, and what data they're accessing

Engineering teams are forced to build bespoke integrations and security controls for each AI use case, leading to long delivery cycles, inconsistent security posture, and uncontrolled risk.

The cost of inaction

Organizations attempting to enable AI agents without standardized architecture experience:

- 6-12 month delivery cycles for each AI use case due to custom development requirements
- 70-90% of engineering capacity consumed by integration plumbing rather than product innovation
- Fragmented security posture with inconsistent controls across AI workflows creating compliance gaps
- Ungoverned MCP servers becoming backdoors into enterprise systems without security validation
- Developer frustration from repetitive integration work preventing focus on strategic initiatives
- Failed AI pilots unable to demonstrate business value due to technical complexity and security constraints

Without a standardized architecture and secure deployment approach, AI integrations often become ungoverned backdoors into enterprise systems.

6-12 month

delivery cycles for each AI use case due to custom development requirements

70-90%

of engineering capacity consumed by integration plumbing rather than product innovation

02

Why now

Enterprises are racing to operationalize agentic AI to improve productivity and customer experience. However, three converging forces make immediate action critical:

Security risk escalation

Insecure MCP servers and ungoverned AI workflows represent a rapidly growing attack surface. Developer-created MCP servers without security review provide direct access to production databases, CRM systems, and financial applications. AI agents with excessive permissions can exfiltrate sensitive data through tool calls. Business logic abuse through autonomous workflows can bypass controls designed for human users.

Organizations cannot afford the risk of ungoverned AI-to-application integrations.

Regulatory compliance requirements

Organizations must comply with emerging AI governance frameworks that require controlled access, auditability, and policy enforcement across AI interactions. The EU AI Act mandates transparency, human oversight, and record-keeping for high-risk AI systems. ISO 42001 requires documented AI management systems with governance controls. NIST AI RMF establishes requirements for managing AI risks including security and privacy.

Regulatory examinations will demand evidence of AI governance—organizations without proper architecture will face findings and potential fines.

Competitive pressure

Organizations that delay standardized AI gateway deployment fall behind competitors who have established secure, scalable AI integration architecture. Fragmented approaches lead to:

- Developer-driven workarounds creating technical debt and security gaps
- Delayed time-to-market for AI capabilities while competitors deploy rapidly
- Uncontrolled exposure of enterprise data through ad-hoc integrations
- Inability to scale AI deployments beyond pilot due to architectural limitations

Organizations must immediately establish an AI-ready integration architecture that embeds security and governance by design.

03

Solution overview

Gruve's **Design and Deployment for AI-Accelerated Services** enables organizations to rapidly convert existing APIs and applications into secure, MCP-compatible tools for enterprise AI agents.

This service designs and deploys a production-ready AI gateway architecture that integrates directly with existing identity systems and application environments. Using automated accelerators, Gruve transforms OpenAPI and Swagger specifications into governed MCP endpoints—allowing AI agents such as enterprise copilots and workflow agents to interact securely with business systems without custom code.

Two-tier engagement model

Service tier	Description	Timeline	Investment
Tier 1: Discovery & architecture design	Comprehensive assessment of existing APIs, identity infrastructure, network topology and data flows. Delivery of secure AI-ready architecture blueprint, Trusted MCP Registry design, and AI Readiness Roadmap.	2-3 weeks	$75,000 - $125,000
Tier 2: Secure deployment & managed integration	End-to-end implementation of AI gateway platform. Conversion of 10-20 priority API endpoints into MCP-compatible tools. OAuth 2.0 and enterprise IAM integration. Context-aware guardrail configuration, session binding protection, PII masking, and telemetry integration with SOC platforms.	3-4 weeks	$185,000 - $325,000
Complete design & deployment package	Combined engagement delivering both architecture design and production deployment in streamlined execution.	5-7 weeks	$245,000 - $425,000

04

Service tier 1: Discovery & architecture design

Objectives

Design secure, production-ready AI gateway architecture tailored to your environment, applications, and security requirements. Deliver comprehensive blueprint enabling rapid deployment of governed AI-to-application capabilities.

Core activities

Week 1: Assessment & discovery

- **API landscape discovery:** Comprehensive inventory and classification of internal, external, and third-party APIs. Identification of APIs prioritized for AI agent enablement. Review of OpenAPI/Swagger documentation completeness and quality. Assessment of API complexity, dependencies, and data sensitivity.
- **Authentication & authorization architecture review:** Analysis of existing IAM infrastructure (OAuth 2.0, SAML, Active Directory). Review of current authentication and authorization patterns. Identification of service accounts, API keys, and credential management approaches. Assessment of role-based access control and permission models.
- **Network & deployment model assessment:** Review of network topology and security zones. Assessment of cloud infrastructure (AWS, Azure, GCP) or hybrid environment. Evaluation of deployment options (SaaS vs. private cloud vs. hybrid). Analysis of connectivity requirements, bandwidth, and latency considerations.
- **Data sensitivity & compliance analysis:** Identification of sensitive data and PII requiring protection. Review of regulatory requirements (GDPR, HIPAA, SOC 2, EU AI Act). Assessment of data residency and sovereignty requirements. Analysis of audit and logging requirements for compliance.
- **Stakeholder interviews:** Security team discussions on threat model and security requirements. Development team discussions on API architecture and dependencies. Compliance team discussions on regulatory requirements. Business stakeholder discussions on AI use cases and priorities.

Week 2: Architecture design

- **AI gateway architecture design:** Selection of optimal deployment model (SaaS, private cloud, hybrid). Design of gateway infrastructure (compute, storage, networking). Architecture for high availability, scalability, and performance. Integration architecture with existing security and operational tools.
- **Trusted MCP registry design:** Design of centralized MCP server registry and governance model. Versioning and lifecycle management approach. Discovery and distribution mechanisms for developers. Approval workflows and security vetting processes.
- **Identity integration design:** OAuth 2.0 and enterprise IAM integration architecture. Authentication flow design for AI agents and applications. Authorization model and role-based access control. Service account and credential management approach.
- **Security & guardrail design:** Context-aware guardrail policies for data protection. Session binding protection architecture. PII masking and data control policies. Rate limiting and abuse prevention mechanisms. Threat detection and monitoring architecture.
- **Observability & telemetry design:** Telemetry architecture for AI workflow monitoring. Integration with SOC platforms and SIEM tools. Logging and audit trail design for compliance. Dashboards and metrics for operational visibility.

Week 3: Roadmap & handoff

- **AI readiness roadmap development:** Phased deployment plan with milestones and timelines. Prioritization of APIs for MCP enablement. Resource requirements and team allocation. Risk assessment and mitigation strategies. Success metrics and measurement approach.
- **Business logic boundaries definition:** Identification of permitted agent actions and workflows. Definition of prohibited operations and safeguards. Approval requirements for sensitive operations. Escalation paths for anomalous behavior.
- **Implementation planning:** Detailed project plan for Tier 2 deployment (if proceeding). Tool procurement requirements and vendor coordination. Infrastructure provisioning requirements. Team training and change management planning.
- **Executive presentation:** Architecture overview for executive stakeholders. Business value and ROI projection. Risk assessment and mitigation approach. Investment requirements and timeline. Go/no-go decision discussion.

Key deliverables

- **AI gateway architecture blueprint:** Detailed technical design documentation including deployment architecture diagrams, component specifications, integration patterns, security controls, and scalability design
- **Trusted MCP registry design specification:** Complete design for governed MCP server registry including approval workflows, versioning strategy, developer access model, and governance policies
- **AI readiness roadmap:** Comprehensive deployment plan including phased implementation timeline, API prioritization, resource allocation, risk mitigation strategies, and success metrics
- **Integration architecture diagrams:** Visual documentation of all integrations including IAM, security tools, applications, and observability platforms
- **Security & guardrail policy framework:** Detailed policies for data protection, access control, session binding, PII masking, and threat detection
- **Executive presentation deck:** Business-focused presentation summarizing architecture, value proposition, risk mitigation, and investment requirements

Typical use cases

Legacy modernization
Organizations with extensive REST API catalogs that need to be made "agent-ready" without manual code refactoring. The design provides blueprint for converting hundreds of existing APIs into secure MCP endpoints using automated accelerators.

Compliance pre-flight
Regulated industries (Healthcare, Finance) requiring a validated security architecture before approving generative AI pilots. The design demonstrates how AI interactions will comply with GDPR, HIPAA, or sector-specific regulations.

Pilot validation
Organizations wanting to validate AI gateway architecture through proof-of-concept before committing to full deployment. Tier 1 delivers production-ready design that can be piloted with limited scope before scaling.

Key outcome

You receive an **AI Readiness Roadmap** and **Trusted MCP Registry design**, documenting exactly how your APIs will be exposed to agents with verified security guardrails. This eliminates the risk of insecure AI interactions with your applications and provides the technical blueprint for immediate deployment.

05

Service Tier 2: Secure deployment & managed integration

Objectives

Execute end-to-end implementation of AI gateway platform transforming APIs into secure, production-ready MCP tools. Deliver operational capability with zero-trust security, comprehensive observability, and validated performance.

Core activities

Week 1: Infrastructure deployment

- **Gateway platform deployment:** Infrastructure provisioning (cloud or on-premises based on design). AI gateway software installation and configuration. High availability and failover configuration. Performance tuning and capacity validation. Security hardening and vulnerability scanning.
- **Network & connectivity:** VPN or secure connectivity setup between gateway and systems. Network security group and firewall configuration. Load balancer and traffic routing setup. DNS and certificate configuration.
- **Development environment:** Separate dev/test environment for MCP development and testing. Version control and change management integration. Testing infrastructure and validation tools.

Week 2: MCP tool creation & integration

- **API to MCP conversion:** Automated conversion of OpenAPI/Swagger specifications to MCP endpoints using accelerators (10-20 APIs based on priority). Manual configuration of complex APIs without complete specifications. Authentication and authorization mapping. Data transformation and enrichment logic. Error handling and retry logic.
- **OAuth 2.0 & IAM integration:** OAuth 2.0 client configuration and credential management. Integration with enterprise IAM for authentication. Role-based access control mapping. Service account provisioning and management. Token validation and refresh configuration.
- **Trusted MCP registry setup:** Registry platform deployment and configuration. MCP server publishing and approval workflows. Versioning and lifecycle management. Developer access and discovery mechanisms.

Week 3: Guardrails & security configuration

- **Context-aware guardrail configuration:** Implementation of data protection policies based on security requirements. PII detection and masking rules. Sensitive operation restrictions and approval requirements. Business logic boundary enforcement. Rate limiting and throttling policies.
- **Session binding protection:** Session-to-IP address binding configuration. Token theft and replay prevention. Session timeout and refresh policies. Multi-factor authentication integration (if required).
- **Data control policies:** Field-level access control based on roles. Dynamic data masking for sensitive information. Audit logging for data access. Encryption at rest and in transit validation.
- **Threat detection configuration:** Anomalous behavior detection rules. Business logic abuse detection. Automated alert generation and escalation. Integration with existing threat detection tools.

Week 4: Observability & validation

- **Telemetry integration:** Real-time telemetry streaming to SOC platforms. Integration with SIEM tools for centralized logging. Metrics collection and dashboard creation. Alert routing and escalation configuration.
- **Comprehensive testing:** Functional testing of all MCP endpoints. Security testing (penetration testing, vulnerability scanning). Performance testing under load. Integration testing with sample AI agents. Edge case and error scenario validation.
- **Knowledge transfer:** Hands-on training for platform administrators. Developer training on MCP consumption from Trusted Registry. Security team training on monitoring and threat response. Documentation handoff (runbooks, troubleshooting guides, configuration details).
- **Operational handoff:** Production readiness review and approval. Transition to operations team. Ongoing support procedures and escalation. 30-day post-deployment support period.

Key deliverables

- **Operational AI gateway infrastructure:** Production-ready platform with high availability, performance tuning, and security hardening
- **10-20 production-ready MCP endpoints:** Converted from priority APIs with authentication, authorization, error handling, and documentation
- **Operational trusted MCP registry:** Deployed registry with governance workflows, developer access, and versioning
- **OAuth/IAM integration:** Complete integration with enterprise identity systems for zero-trust authentication
- **Guardrail policies implemented:** Context-aware security policies, session binding, PII masking, and data controls operational and tested
- **Telemetry integration with SOC platforms:** Real-time streaming to SIEM/SOAR with dashboards and alerting
- **Comprehensive test results and validation report:** Documentation of all testing with results, identified issues, and resolutions
- **Training completed for all teams:** Platform administrators, developers, and security teams trained and certified
- **Operational runbooks and documentation:** Complete operational documentation including troubleshooting guides, escalation procedures, and configuration details

Typical use cases

Agentic productivity
Enabling an internal AI agent to safely perform complex tasks like querying SQL databases, updating CRM records, or managing helpdesk tickets. The deployment provides secure MCP endpoints for all required systems with proper authentication and data controls.

Autonomous customer experience
Deploying a secure shopping companion or product advisor that interacts directly with inventory and checkout APIs. The implementation ensures customer data protection while enabling rich agent capabilities.

Multi-system workflow
Automation Enabling AI agents to orchestrate workflows across multiple enterprise systems (ERP, CRM, ticketing, documentation) with consistent security and governance.

Key outcome

Your organization achieves a **Production-Ready AI Gateway** that scales securely across all departments. By centralizing governance, you reduce AI integration costs by up to 70% while ensuring that all agentic interactions remain visible, compliant, and protected against evolving threats.

06

Benefits of Gruve's design & deployment service

No-code AI enablement: Rapidly transform internal, external, or SaaS applications into agent-ready tools without custom development. Automated accelerators convert OpenAPI/Swagger specifications into MCP-compatible endpoints in hours vs. weeks of manual coding. Organizations achieve 70-90% reduction in engineering effort compared to custom MCP development for each integration.

Zero-trust AI access: AI tool calls are protected by enterprise identity and authorization controls. Seamless integration with existing OAuth 2.0 and IAM infrastructure ensures agents only access data they're explicitly authorized to use, maintaining consistent authentication standards across all AI interactions. Prevents unauthorized data access and lateral movement by AI agents.

Trusted MCP registry: Centralized, governed registry of vetted MCP servers prevents use of insecure or unverified tools. Eliminates risk of developers downloading untrusted servers from the internet or creating ungoverned custom servers. Establishes governance and security standards for all AI-to-application integrations.

Rapid time-to-value: AI-ready endpoints delivered in 5-7 weeks rather than 6-12 months of custom development. MVP prototypes in 2-3 weeks provide early validation, with production-ready capabilities following in 3-4 weeks. Organizations achieve 4-5× ROI through accelerated time-to-market and elimination of custom integration development.

Production-grade scalability: Supports SaaS and private cloud deployments across major cloud providers (AWS EKS, Azure AKS, GCP GKE). Architecture designed for high availability, performance under load, geographic distribution, and disaster recovery. Eliminates common pitfall of POC architectures failing at production scale.

End-to-end observability: Real-time telemetry of AI workflows and tool calls for threat detection and auditing. Streaming telemetry integrates with SOC platforms (Splunk, Microsoft Sentinel, QRadar) providing comprehensive visibility into agent behavior, business logic abuse detection, and compliance audit trails.

Security-first architecture: Security and governance controls embedded at gateway and protocol level rather than bolted on later. Context-aware guardrails, session binding protection, PII masking, and data control policies prevent sensitive data exposure and unauthorized access from design phase through deployment.

Accelerated MCP enablement: Automated accelerators convert API definitions into MCP-compatible tools with minimal engineering effort. Proven conversion methodology handles authentication, authorization, data mapping, error handling, and documentation automatically. Production-ready MCP endpoints delivered rapidly without extensive development cycles.

Enterprise identity integration: AI interactions fully aligned with existing OAuth and IAM platforms. Maintains enterprise authentication standards, role-based access control, and authorization policies without requiring separate identity infrastructure for AI agents.

Scalable foundation for 100+ application integrations: Scalable architecture supporting integration with diverse systems including CRM platforms (Salesforce, Microsoft Dynamics), databases (PostgreSQL, MySQL, SQL Server), ticketing systems (ServiceNow, Jira), ERP applications (SAP, Oracle), SaaS tools, and custom applications. Single gateway provides consistent security and governance across all integrations.

07

What makes Gruve's design & deployment unique

We architect *and* deploy what others only sell

Unique advantage of same team designing and deploying ensures design fidelity without "lost in translation" between teams. Architectural decisions informed by implementation realities, not just theoretical design. Continuous learning loop improving both design and deployment methodologies. **End-to-end accountability** for outcomes from architecture through operational capability.

Automated accelerators for rapid MCP enablement

Proprietary accelerators convert OpenAPI/Swagger specifications into MCP-compatible endpoints in hours vs. weeks of custom development. Proven conversion methodology handles authentication, authorization, data mapping, error handling automatically. Organizations achieve **70-90% reduction in engineering effort** compared to manual MCP development.

Security-first, not bolted-on later

Security and governance embedded at gateway and protocol level from design phase, not added later. Context-aware guardrails, zero-trust authentication, session binding, PII masking, and threat detection built into architecture. **Eliminates security gaps** common in DIY implementations where security becomes afterthought.

Vendor-agnostic integration expertise

Deep experience integrating with all major enterprise platforms: IAM systems (Azure AD, Okta, Ping, ForgeRock), cloud infrastructure (AWS, Azure, GCP), SOC platforms (Splunk, Microsoft Sentinel, QRadar), and enterprise applications (Salesforce, ServiceNow, SAP). Independence ensures **best-of-breed architecture** without vendor lock-in or bias.

Production-grade scalability from day one

Architecture designed for enterprise scale, not pilot/POC. Supports high availability, performance under load, geographic distribution, and disaster recovery. Deployment patterns validated across SaaS, Kubernetes-based private cloud, and hybrid environments. **Eliminates common pitfall** of POC architectures failing at production scale requiring expensive rework.

Trusted MCP registry governance

Unique approach establishing centralized, governed registry preventing insecure developer-created servers. Combines convenience of easy MCP consumption with governance and security vetting. **Eliminates major risk** of developers downloading untrusted tools from internet or creating ungoverned custom servers.

Comprehensive observability and threat detection

Real-time telemetry providing complete visibility into AI agent behavior and tool usage. Integration with SOC platforms enabling detection of business logic abuse, sensitive data exfiltration, and anomalous patterns. **Compliance audit trails built-in** from deployment rather than added later.

Proven methodology across diverse industries

Battle-tested approach refined across financial services, healthcare, retail, manufacturing, and technology customers. Experience with diverse requirements (GDPR, HIPAA, SOC 2, PCI-DSS, EU AI Act). Knowledge of industry-specific challenges and best practices **accelerates deployment and reduces risk.**

Knowledge transfer, not dependency

Goal is your independence, not consulting dependency. Hands-on training throughout deployment (not just at end). Comprehensive documentation enabling self-sufficiency. Operational runbooks for ongoing management. **Gradual handoff** maintaining support during transition ensuring you own the platform post-deployment.

4-5× ROI through engineering efficiency

Organizations realize measurable ROI through: elimination of 70-90% of engineering effort for AI enablement, acceleration from 6-12 months DIY to 5-7 weeks professional deployment, avoidance of technical debt requiring expensive remediation, and scalable foundation supporting 100+ application integrations without rework. **Our fee typically represents 1-2 months of engineering cost** while saving 6-9 months of effort.

08

Common deployment use cases

Enabling internal AI agents to query enterprise databases securely
AI copilots need to access customer data, order history, inventory information, and transactional records from SQL databases. The AI gateway provides secure MCP endpoints with:

- Row-level security based on agent identity and user context
- Query result filtering preventing access to restricted data
- Audit logging of all database queries for compliance
- Rate limiting preventing excessive database load

Result: AI agents can query databases securely while protecting sensitive data and maintaining performance.

Allowing AI copilots to update CRM or ticketing systems
Customer service AI agents need to create tickets, update customer records, log interactions, and escalate issues in ServiceNow or Salesforce. The implementation provides:

- Field-level permissions controlling what agents can read/write
- Approval workflows for sensitive operations (account changes, refunds)
- Data validation preventing invalid updates
- Integration with existing approval and audit processes

Result: AI agents can perform business operations with appropriate controls and governance.

Deploying autonomous shopping or advisory agents
E-commerce AI agents help customers find products, check availability, apply discounts, and complete purchases. The deployment ensures:

- Real-time inventory access with proper caching
- Secure payment processing integration
- Fraud detection and prevention
- Customer data protection and PII masking

Result: Customers receive AI-powered assistance while maintaining security and compliance.

Enabling AI-driven workflow automation across SaaS platforms
Business process AI agents orchestrate workflows across multiple systems (create lead in Salesforce → create ticket in Jira → send notification in Slack → update dashboard). The architecture provides:

- Consistent authentication across all platforms
- Error handling and retry logic for reliability
- Workflow observability and audit trails
- Centralized governance for cross-platform workflows

Result: Complex multi-system workflows automated with security and reliability.

09

Expected outcomes

Tier 1: Discovery & architecture design outcomes

- **Validated architecture blueprint:** Production-ready design reviewed and approved by security, compliance, and engineering stakeholders
- **Clear AI readiness roadmap:** Phased deployment plan with specific milestones, resource requirements, and success criteria
- **Risk mitigation strategy:** Identified risks with specific mitigation approaches preventing deployment failures
- **Stakeholder alignment:** Executive, security, development, and business stakeholders aligned on approach and investment
- **Go/No-go decision confidence:** Clear understanding of what deployment will deliver enabling informed investment decision
- **Foundation for rapid deployment:** Detailed design enabling immediate Tier 2 execution (if approved) without additional planning cycles

Complete package outcomes

All Tier 1 and Tier 2 outcomes delivered in streamlined 5-7 week engagement with consistent team and seamless execution from design through deployment.

Tier 2: Secure deployment & managed integration outcomes

- **Operational AI gateway:** Production-ready platform serving live AI agent traffic with validated performance and security
- **10-20 secure MCP endpoints:** Priority APIs converted to agent-ready tools with authentication, authorization, and observability
- **Zero-trust authentication operational:** Enterprise IAM integration ensuring only authorized agents access appropriate data
- **Security controls validated:** Guardrails, session binding, PII masking tested and operational preventing data exposure
- **Full observability:** Real-time telemetry streaming to SOC platforms providing visibility into all AI interactions
- **70% reduction in integration costs:** Elimination of custom development for each AI use case through reusable architecture
- **4-5× ROI:** Measurable return through engineering efficiency, accelerated time-to-market, and avoided technical debt
- **Team capability:** Platform administrators, developers, and security teams trained and capable of ongoing operations

10

Investment and engagement model

Tier 1: Discovery & architecture design

Service fee: $75,000 - $125,000

Pricing factors:
- Organization size and complexity (number of APIs, systems, geographic distribution)
- API infrastructure maturity (quality of existing documentation and specifications)
- Integration complexity (number of IAM systems, security tools requiring integration)
- Compliance requirements (number of regulations, audit requirements)
- Timeline acceleration (compressed schedule requires additional resources)

Payment structure:
- 50% at engagement kickoff
- 50% at final deliverable acceptance

Timeline: 2-3 weeks from kickoff to final presentation

Post-engagement support: 30 days of email/phone support for questions on design

Tier 2: Secure deployment & managed integration

Service fee: $185,000 - $325,000

Pricing factors:
- Number of MCP endpoints (10-20 standard, additional endpoints extra)
- Infrastructure complexity (SaaS vs. private cloud vs. on-premises)
- Integration scope (number of systems requiring integration)
- Customization requirements (custom guardrails, unique workflows)
- Geographic distribution (multi-region deployments)
- Timeline acceleration (compressed schedule requires additional resources)

Payment structure:
- 40% at engagement kickoff
- 30% at infrastructure deployment completion
- 30% at final acceptance and handoff

Timeline: 3-4 weeks from kickoff to operational handoff

Post-deployment support: 30 days included (email, phone, remote troubleshooting)

Complete design & deployment package

Service fee: $245,000 - $425,000

Pricing factors:
Combination of factors from both tiers plus:

- Integrated execution providing efficiency gains vs. separate engagements
- Reduced overhead from single-team continuity
- Streamlined governance and approvals

Payment structure:

- 430% at engagement kickoff
- 30% at design completion
- 20% at infrastructure deployment
- 20% at final acceptance and handoff

Timeline: 5-7 weeks end-to-end

Post-deployment support: 30 days included

What's included

All Tier 1 engagements:

- Comprehensive API assessment and classification
- Authentication/authorization architecture review
- Security and compliance requirements analysis
- AI gateway architecture blueprint
- Trusted MCP Registry design
- AI Readiness Roadmap
- Stakeholder presentations
- 30 days post-design support

All Tier 2 engagements:

- Complete AI gateway infrastructure deployment
- 10-20 production-ready MCP endpoints
- OAuth 2.0 and enterprise IAM integration
- Context-aware guardrail configuration
- Session binding and PII masking
- Telemetry integration with SOC platforms
- Comprehensive testing and validation
- Knowledge transfer and training
- Operational documentation and runbooks
- 30 days post-deployment support

What's not included (client responsibility)

- AI gateway software licensing (Cequence or alternative)
- Cloud infrastructure costs (compute, storage, networking)
- Third-party tool licensing (IAM, SIEM, SOAR)
- Internal team time commitment (estimated 200-400 hours)
- Travel expenses if onsite presence required (estimated separately)
- Ongoing operations and platform management after handoff
- Additional MCP endpoints beyond standard scope (available as add-on)

Optional add-on services

- **Accelerated timeline:** Compressed delivery schedule (15-25% premium on base fee)
- **Extended MCP endpoint catalog:** Additional endpoints beyond standard scope ($8K-$15K per API)
- **API documentation services:** Creating OpenAPI specs for undocumented APIs ($25K-$45K)
- **Extended post-deployment support:** Continued support beyond 30 days ($5K-$10K/month)
- **Managed services:** Ongoing 24/7 platform operations and optimization (custom pricing based on scope)
- **Training programs:** Advanced training for additional team members ($5K-$10K per program)

11

Getting started

Prerequisites for engagement

- Executive sponsorship and approved budget
- Existing APIs requiring AI agent enablement
- Enterprise IAM infrastructure (OAuth 2.0, SAML, or similar)
- Cloud infrastructure or hosting environment
- Stakeholder availability for collaborative design (50-100 hours Tier 1, 200-400 hours Tier 2)
- Clear AI agent use cases and deployment timeline

Engagement process

Step 1: Initial consultation
60-90 minutes

- Review AI agent deployment plans and use cases
- Discuss API infrastructure and documentation maturity
- Assess security and governance requirements
- Review timeline and business drivers
- Align on engagement approach (Tier 1, Tier 2, or Complete Package)
- Provide detailed proposal and investment estimate

Step 2: Engagement planning
1 week

- Finalize statement of work and scope
- Schedule assessment activities and stakeholder interviews
- Establish governance and communication cadence
- Mobilize Gruve team
- Complete prerequisites and access provisioning

Step 3: Design phase - Tier 1
2-3 weeks

- Conduct comprehensive assessment
- Design AI gateway architecture
- Develop Trusted MCP Registry design
- Create AI Readiness Roadmap
- Present to stakeholders and gain approval

Step 4: Deployment phase - Tier 2
3-4 weeks

- Deploy AI gateway infrastructure
- Create MCP endpoints from priority APIs
- Configure security guardrails and integrations
- Validate through comprehensive testing
- Transfer knowledge and handoff to operations

Step 5: Post-deployment support
30 days

- Email and phone support for operational questions
- Monthly check-in on performance and optimization
- Guidance on expansion or enhancement
- Transition to extended support or independent operation

12

Accelerate secure AI integration

Engage Gruve's Design and Deployment service to establish a secure, governed AI gateway and MCP infrastructure.

Schedule a 60-90 minute consultation to review your AI agent deployment plans, discuss API infrastructure maturity, assess security requirements, and align on engagement approach.

Time to value

- Design phase: 2-3 weeks to production-ready architecture blueprint
- Deployment phase: 3-4 weeks to operational AI gateway
- Complete package: 5-7 weeks from kickoff to production deployment

Fixed-price engagement with clear deliverables

- Tier 1: $75K-$125K for comprehensive design
- Tier 2: $185K-$325K for production deployment
- Complete package: $245K-$425K for end-to-end solution

Expected ROI
Organizations realize 4-5× ROI by eliminating extensive custom coding, reducing engineering time 70-90%, and accelerating agentic AI project delivery from months to weeks.

This engagement enables your organization to rapidly and safely deploy enterprise AI agents while maintaining full security, visibility, and compliance from day one.

Website: https://www.gruve.ai/

Email: info@gruve.ai

Request consultation:
https://www.gruve.ai/contact

Prerequisites:

- AI agent deployment plans
- Existing API infrastructure
- Approved budget for AI enablement

This solution brief is current as of February 2026. For most up-to-date service offerings, pricing, and methodology, contact Gruve.ai directly.