

AI Control Tower

Security & Compliance Whitepaper

September 12, 2026

AI Control Tower

Security & Compliance Whitepaper

Prepared for: AI Control Tower

Date: September 12, 2026

Version: 2.0

Executive Summary

AI Control Tower is an enterprise-grade AI governance platform that provides comprehensive monitoring, compliance, and control over artificial intelligence systems. This whitepaper outlines our security architecture, compliance frameworks, and data protection measures.

Key Security Features

- End-to-End Encryption:** TLS 1.3 for data in transit, AES-256 for data at rest
- Multi-Tenant Isolation:** Row-Level Security (RLS) ensures complete data separation
- Zero-Trust Architecture:** All requests authenticated and authorized
- Audit Logging:** Comprehensive activity logs for all user and system actions
- Anomaly Detection:** Real-time detection of security threats and data breaches
- Incident Response:** Automated alerting and integration with enterprise security tools

1. Security Architecture

1.1 Infrastructure Security

Cloud Infrastructure:

- Hosted on enterprise-grade cloud platforms (Vercel, Supabase)
- Multi-region deployment for redundancy
- Automated failover and disaster recovery
- 99.9% uptime SLA

Network Security:

- HTTPS/TLS 1.3 enforced for all connections
- DDoS protection via CDN (Cloudflare/Vercel)
- Web Application Firewall (WAF) enabled
- IP allowlisting available for enterprise customers

Database Security:

- PostgreSQL with Row-Level Security (RLS)
- Encrypted at rest with AES-256
- Automated daily backups with point-in-time recovery
- Database connection pooling with encrypted credentials

1.2 Application Security

Authentication & Authorization:

- Multi-factor authentication (MFA) support
- OAuth 2.0 / OIDC integration
- Role-Based Access Control (RBAC)
- API key authentication with SHA-256 hashing
- Session management with secure cookies
- Automatic session expiration

Access Control:

- Principle of least privilege enforced
- Just-in-time access for administrative operations
- Quarterly access reviews
- Automated deprovisioning on account termination

Data Protection:

- Input validation on all API endpoints
- SQL injection prevention via parameterized queries

- XSS protection with Content Security Policy (CSP)
- CSRF protection with anti-CSRF tokens
- Rate limiting to prevent abuse
- PII detection system with automatic alerts

1.3 Encryption

Data in Transit:

- TLS 1.3 with strong cipher suites
- Certificate pinning for mobile applications
- Encrypted webhooks with HMAC-SHA256 signatures
- Encrypted API communications

Data at Rest:

- AES-256 encryption for database
 - SHA-256 hashing for API keys (irreversible)
 - Encrypted backup storage
 - Secure key management via cloud provider KMS
-

2. Compliance Frameworks

2.1 SOC 2 Type II Compliance

AI Control Tower is designed to meet SOC 2 Trust Service Criteria across all five categories:

Security (Common Criteria):

- CC1: Control Environment - Code of conduct, ethics policies
- CC2: Communication and Information - Internal security communications
- CC3: Risk Assessment - Annual risk assessments conducted
- CC4: Monitoring Activities - Continuous security monitoring
- CC5: Control Activities - Security controls implemented and tested
- CC6: Logical and Physical Access Controls - MFA, RLS, encryption
- CC7: System Operations - 24/7 monitoring, incident response
- CC8: Change Management - Version control, testing, approvals
- CC9: Risk Mitigation - Vulnerability scanning, penetration testing

Availability:

- A1.1: 99.9% uptime SLA maintained

- A1.2: Multi-region redundancy
- A1.3: Automated failover and disaster recovery

Confidentiality:

- C1.1: Multi-tenant data isolation via RLS
- C1.2: Secure data disposal on account deletion
- C1.3: PII detection and protection

Processing Integrity:

- PI1.1: Input validation and data integrity checks
- PI1.2: Error detection and correction procedures
- PI1.3: Audit trails for all data processing

Privacy:

- P1.1: Privacy policy and data usage disclosure
- P2.1: Data subject rights (access, correction, deletion)
- P3.1: Data retention and disposal policies
- P4.1: Incident response and breach notification

2.2 ISO 27001 Information Security Management

Control Objectives Implemented:**A.5 - Information Security Policies:**

- Documented security policies and procedures
- Regular policy reviews and updates
- Employee security awareness training

A.6 - Organization of Information Security:

- Defined security roles and responsibilities
- Segregation of duties
- Contact with authorities and special interest groups

A.9 - Access Control:

- Access control policy
- User access management
- User responsibilities
- System and application access control

A.10 - Cryptography:

- Cryptographic controls (TLS 1.3, AES-256)
- Key management procedures

A.12 - Operations Security:

- Change management procedures
- Capacity management
- Malware protection
- Backup procedures
- Logging and monitoring

A.14 - System Acquisition, Development and Maintenance:

- Secure development lifecycle
- Security requirements analysis
- Secure coding practices
- Security testing

A.16 - Information Security Incident Management:

- Incident response procedures
- Evidence collection
- Lessons learned process

A.18 - Compliance:

- Legal and regulatory compliance
- Information security reviews
- Technical compliance reviews

2.3 GDPR Compliance

AI Control Tower provides GDPR-compliant data processing:

Lawfulness of Processing (Article 6):

- Legitimate interest basis for AI monitoring
- Consent mechanisms where required
- Contractual necessity for service delivery

Data Subject Rights:

- ****Right of Access (Article 15):**** Self-service data export
- ****Right to Rectification (Article 16):**** Profile editing
- ****Right to Erasure (Article 17):**** Account deletion

- **Right to Data Portability (Article 20):** JSON/CSV export
- **Right to Object (Article 21):** Opt-out mechanisms

Data Protection by Design and Default (Article 25):

- Privacy-first architecture
- Minimal data collection
- Pseudonymization where possible
- Encryption by default

Security of Processing (Article 32):

- State-of-the-art encryption
- Regular security testing
- Pseudonymization and encryption
- Ability to restore availability

Data Breach Notification (Article 33-34):

- 72-hour breach notification procedures
- Automated breach detection via anomaly detection
- Communication templates for data subjects

2.4 HIPAA Compliance

For healthcare organizations, AI Control Tower supports HIPAA compliance:

Administrative Safeguards:

- Security Management Process
- Assigned Security Responsibility
- Workforce Security (background checks, training)
- Information Access Management (role-based access)
- Security Awareness and Training
- Security Incident Procedures
- Contingency Plan (disaster recovery)

Physical Safeguards:

- Facility Access Controls (cloud data center security)
- Workstation Security
- Device and Media Controls

Technical Safeguards:

- Access Control (unique user IDs, emergency access, encryption)
- Audit Controls (comprehensive logging)
- Integrity Controls (data validation, checksums)
- Person or Entity Authentication (MFA)
- Transmission Security (TLS 1.3)

PHI Protection:

- Automatic PHI detection in AI decisions
- Alert on PHI exposure
- Encrypted storage and transmission
- Minimum necessary access principle

3. Data Protection

3.1 Data Classification

****Public:**** Marketing materials, public documentation

****Internal:**** Product roadmaps, internal communications

****Confidential:**** Customer data, AI decisions, analytics

****Restricted:**** API keys, passwords, encryption keys

3.2 Data Retention

- ****AI Decisions:**** Retained according to customer plan (30 days to unlimited)
- ****Audit Logs:**** Retained for 1 year minimum
- ****Backups:**** 30-day retention for point-in-time recovery
- ****User Data:**** Deleted within 30 days of account closure

3.3 Data Disposal

- ****Secure Deletion:**** Multi-pass overwrite for sensitive data
- ****Database Deletion:**** Hard delete from production and backups
- ****Backup Expiry:**** Automated deletion after retention period
- ****Media Destruction:**** Physical media destroyed per NIST 800-88

4. Incident Response

4.1 Security Incident Classification

Severity Levels:

- **P1 (Critical):** Data breach, system compromise
- **P2 (High):** Unauthorized access, service disruption
- **P3 (Medium):** Policy violation, anomaly detection
- **P4 (Low):** Minor security event

4.2 Incident Response Process

1. **Detection:** Automated anomaly detection, user reports
2. **Containment:** Isolate affected systems, revoke access
3. **Investigation:** Root cause analysis, forensics
4. **Eradication:** Remove threat, patch vulnerabilities
5. **Recovery:** Restore services, verify integrity
6. **Lessons Learned:** Post-incident review, improvements

4.3 Notification Procedures

- **Internal:** Security team notified within 15 minutes
 - **Customer:** Notification within 24 hours for P1/P2
 - **Regulatory:** GDPR 72-hour notification if applicable
 - **Public:** Public disclosure if required by law
-

5. Third-Party Security

5.1 Vendor Management

Vendor Selection:

- SOC 2 Type II audit reports required
- Security questionnaire completion
- Contract security terms negotiation

Key Vendors:

- **Vercel:** Cloud hosting platform (SOC 2, ISO 27001)
- **Supabase:** Database platform (SOC 2, ISO 27001, HIPAA)
- **Stripe:** Payment processing (PCI DSS Level 1)

5.2 Data Processing Agreements

- Standard Contractual Clauses (SCCs) for EU data
 - Data Processing Agreements (DPAs) with all vendors
 - HIPAA Business Associate Agreements (BAAs) available
-

6. Monitoring & Auditing

6.1 Security Monitoring

Real-Time Monitoring:

- Anomaly detection (cost spikes, latency, PII)
- Failed authentication attempts
- Suspicious access patterns
- Rate limit violations

Security Tools:

- SIEM integration (Datadog)
- Vulnerability scanning (quarterly)
- Penetration testing (annual)
- Dependency scanning (automated)

6.2 Audit Logging

Events Logged:

- User authentication (success/failure)
- API key creation/revocation
- Data access and modifications
- Configuration changes
- Policy violations
- Security incidents

****Log Retention:**** 1 year minimum

****Log Protection:**** Tamper-evident, encrypted storage

****Log Access:**** Restricted to security team

7. Business Continuity

7.1 Disaster Recovery

Recovery Objectives:

- **RTO (Recovery Time Objective):** 4 hours
- **RPO (Recovery Point Objective):** 1 hour

Backup Strategy:

- Automated daily backups
- Point-in-time recovery
- Multi-region backup storage
- Weekly backup testing

7.2 Incident Response Team

Security Team:

- Chief Security Officer (CSO)
- Security Engineers
- Incident Response Team
- 24/7 on-call rotation

Escalation Path:

- L1: Security Engineer (15 min response)
 - L2: Senior Security Engineer (30 min response)
 - L3: CSO/CTO (1 hour response)
-

8. Security Testing

8.1 Testing Program

Vulnerability Scanning:

- Weekly automated scans
- Dependency vulnerability checks
- Static code analysis

Penetration Testing:

- Annual third-party penetration tests
- Quarterly internal security assessments
- Bug bounty program for responsible disclosure

Security Audits:

- SOC 2 Type II audit (annual)
 - Internal security audits (quarterly)
 - Code reviews for all changes
-

9. Employee Security

9.1 Security Awareness

Training Program:

- Security onboarding for new hires
- Annual security awareness training
- Phishing simulation exercises
- Incident response drills

9.2 Access Management

Background Checks:

- Required for all employees
- Enhanced checks for security team

Access Reviews:

- Quarterly access reviews
 - Just-in-time access for sensitive operations
 - Automated deprovisioning on termination
-

10. Contact Information

Security Team:

- Email: security@aicontroltower.com
- PGP Key: Available on request
- Bug Bounty: security@aicontroltower.com

Incident Reporting:

- 24/7 Hotline: +1-xxx-xxx-xxxx
- Email: incidents@aicontroltower.com

- Response Time: 15 minutes for P1/P2

Compliance Inquiries:

- Email: compliance@aicontroltower.com
 - Audit Coordination: audits@aicontroltower.com
-

Appendix A: Compliance Certifications

- SOC 2 Type II (in progress)
- ISO 27001 (planned Q2 2025)
- GDPR compliant
- HIPAA compliant (BAA available)

Appendix B: Security Controls Matrix

[Detailed mapping of security controls to compliance frameworks]

Appendix C: Incident Response Runbooks

[Step-by-step procedures for common security incidents]

Document Control:

- Version: 2.0
 - Last Updated: September 12, 2026
 - Next Review: 12/11/2026
 - Owner: Chief Security Officer
 - Classification: Confidential
-

This whitepaper is provided for informational purposes and represents AI Control Tower's current security practices. Security measures are continuously improved and updated.