Whiteout AI

Enterprise AI Governance Platform

Security & Compliance Whitepaper

Version 2.0 | February 2026
Groovy Security, Inc.
Classification: Public

Table of Contents

1. Executive Summary

The rapid adoption of generative AI tools in enterprise environments has created an unprecedented governance challenge. Organizations must balance the productivity benefits of AI assistants like ChatGPT, Claude, Gemini, and Copilot with the very real risks of data leakage, regulatory violations, and loss of intellectual property.

Whiteout AI is an enterprise-grade AI governance platform developed by Groovy Security that enables organizations to safely adopt generative AI tools while enforcing compliance, data security, and auditability. Unlike traditional security tools that either block AI entirely or fail to understand conversational data flows, Whiteout AI provides intelligent, real-time interception with contextual compliance evaluation.

Key Differentiators

- **99.5% Compliance Accuracy** — Rigorously validated across comprehensive test suites covering PHI, PII, GDPR, financial data, and more
- **Universal Coverage** — Browser extension (Chrome, Firefox, Edge, Safari), desktop applications (macOS, Windows), and internal secure AI interface
- **Real-Time Interception** — Sub-second policy evaluation with streaming feedback
- **65 Pre-Built Policies** — Comprehensive policy library spanning 9 regulatory categories, immediately deployable
- **Fail-Safe Design** — The system is designed to never block productivity due to technical issues
- **Enterprise Integration** — SSO/IDP support, SIEM integration, and developer tool integration

The Business Case

Organizations deploying Whiteout AI gain:

- **Risk Reduction:** Prevent inadvertent disclosure of PHI, PII, source code, and confidential business information to external AI services
- **Compliance Assurance:** Demonstrate regulatory compliance (HIPAA, GDPR, SOX, PCI-DSS) with complete audit trails of all AI interactions
- **Operational Visibility:** Understand how AI tools are being used across the organization, identify training needs, and quantify productivity gains
- **Controlled Adoption:** Enable AI usage for approved use cases while maintaining security boundaries

2. The Enterprise AI Governance Challenge

The New Threat Landscape

The emergence of large language models and generative AI tools has fundamentally altered the enterprise security landscape. Unlike traditional applications where data flows are predictable and controllable, AI assistants create a new category of risk: conversational data exfiltration.

Employees interact with AI assistants in natural language, often sharing context that would never be entered into a traditional form or application. A developer asking for help debugging code may paste proprietary source code. A healthcare worker seeking to draft a patient letter may include protected health information. A financial analyst requesting a presentation template may share unreleased earnings data.

These interactions are rarely malicious — they're the natural result of productivity-focused employees using powerful tools. But the consequences can be severe:

- **Data Breach Liability:** Information shared with external AI services may be stored, used for training, or inadvertently exposed
- **Regulatory Violations:** HIPAA, GDPR, and industry-specific regulations prohibit sharing certain data categories with unauthorized third parties
- **Intellectual Property Loss:** Source code, algorithms, and trade secrets shared with AI services may lose legal protection
- **Competitive Exposure:** Strategic plans, M&A; details, and unreleased product information could reach competitors

The Governance Gap

Traditional security tools are poorly suited to address conversational AI risks:

DLP (Data Loss Prevention) systems designed for file transfers and email cannot effectively analyze the unstructured, conversational nature of AI interactions. A prompt asking "Can you help me improve this patient discharge process?" contains no keywords that trigger traditional DLP rules, yet may be followed by sensitive details.

CASB (Cloud Access Security Broker) solutions can block access to AI services entirely, but this creates a productivity crisis. Blocking ChatGPT doesn't prevent employees from using it — it pushes usage to personal devices where there is zero visibility.

Training and Policy alone cannot address the challenge. Employees have good intentions but lack the expertise to identify every category of sensitive data in real-time while trying to be productive.

The Productivity Imperative

Despite the risks, blocking AI tools is increasingly untenable. Organizations that restrict AI access face:

- **Competitive Disadvantage:** Competitors leveraging AI gain productivity advantages
- **Talent Retention:** Knowledge workers increasingly expect AI tool access
- **Shadow IT:** Employees circumvent restrictions using personal devices
- **Innovation Stagnation:** AI tools accelerate ideation and problem-solving

The challenge is not whether to allow AI — it's how to enable it safely.

3. Solution Overview

The Whiteout AI Approach

Whiteout AI takes a fundamentally different approach to AI governance: intelligent interception with contextual compliance. Rather than blocking AI tools or relying on post-hoc detection, Whiteout AI evaluates every prompt and file upload against organizational policies in real time — before any data reaches external AI services.

This approach provides:

- **Preventive Control:** Sensitive data is blocked before it reaches external AI services — not detected after the fact
- **Contextual Understanding:** The compliance engine uses AI-powered analysis to understand natural language context, not just pattern matching
- **User Education:** When prompts are blocked, users see exactly which policies were violated and why, creating a feedback loop that improves behavior
- **Zero Productivity Impact:** Compliant prompts proceed with imperceptible latency; users only experience friction when attempting policy violations

Core Capabilities

Multi-Surface Coverage

Whiteout AI provides unified policy enforcement across all AI touchpoints:

- Browser-based AI tools (ChatGPT, Claude, Gemini, Copilot, and more)
- Desktop AI applications (ChatGPT Desktop, Claude Desktop)
- Internal secure AI interface for sensitive queries
- Developer tool integration via MCP (Model Context Protocol)

Real-Time Policy Evaluation

- Sub-second evaluation with streaming feedback
- 65 pre-built policies across 9 regulatory categories
- Request tailored policies for your organization’s specific needs
- Group-based policy assignment for departmental compliance

Complete Audit Trail

- Prompt content, user identity, timestamp, and device logged
- Policy evaluation results and risk scores
- Exportable for compliance reporting (PDF/CSV)

Enterprise Integration

- SSO/IDP support (SAML, OAuth) with automatic organization detection
- SIEM integration for security event correlation
- API access for custom integrations

4. Platform Architecture

System Overview

Whiteout AI is architected as a distributed system with components optimized for their specific deployment contexts. The platform consists of four primary components connected through a secure backend:

Backend Services

The backend is the intelligence center of the platform, providing intelligent compliance evaluation, risk assessment, integration management, and immutable audit logging — all built on modern, high-performance frameworks optimized for low-latency async processing and secure API design.

Admin Dashboard

A cross-platform desktop application providing policy management, analytics, audit log viewing, integration configuration, user management, and a secure internal AI interface. Built with modern frameworks for secure, native desktop deployment with OS-level credential storage.

Browser Extension

The browser extension provides real-time interception for web-based AI tools with full support across Chrome, Firefox, Edge, and Safari.

Capabilities:

- Prompt interception before submission to AI services
- File upload scanning for drag-drop, paste, and file picker
- External tool blocking for controlled tool access
- Coverage heartbeat for compliance reporting

Supported AI Platforms:

- OpenAI (ChatGPT), Anthropic (Claude), Google (Gemini)
- Microsoft (Copilot), Perplexity AI, Mistral, xAI (Grok)
- DeepSeek, Poe, You.com, HuggingFace, and more

Desktop Guard

Desktop Guard provides native application monitoring for desktop AI clients on both macOS and Windows, using native OS-level APIs for deep application integration.

Core Features:

- Keyboard interception in AI applications
- File drop scanning for drag-and-drop operations
- File picker interception for attachment scanning
- Visual monitoring status indicator
- External tool policy enforcement

Supported Applications: ChatGPT Desktop and Claude Desktop on macOS and Windows.

5. The Compliance Engine

Architecture Overview

The Compliance Engine is the heart of Whiteout AI's policy enforcement capability. Unlike traditional DLP systems that rely on pattern matching and keyword detection, the Compliance Engine uses large language model (LLM) analysis to understand the semantic meaning and context of every prompt — achieving dramatically higher accuracy with fewer false positives.

Evaluation Process

- **Policy Resolution:** Based on the user's group membership, the engine loads applicable policies from the policy library
- **Contextual Analysis:** The prompt is analyzed with specific evaluation criteria for each policy
- **Decision Aggregation:** Results from all policies are combined into a compliance decision
- **Risk Scoring:** A numerical risk score is calculated based on policy severity and confidence
- **Audit Logging:** The evaluation result is logged with full context

Contextual Intelligence vs. Pattern Matching

Traditional DLP relies on regular expressions and keyword matching, which produces high false positive rates and fails to understand context. Consider these prompts:

> *"John Smith was diagnosed with diabetes" → PHI violation (patient name + diagnosis)*

> *"How many patients were diagnosed with diabetes this month?" → Compliant (aggregate statistics)*

Pattern matching cannot distinguish between these prompts — both contain "diagnosed" and "diabetes." The Compliance Engine's AI-powered evaluation understands the semantic difference: the first reveals individual patient information while the second requests aggregate statistics. This contextual understanding is what enables 99.5% accuracy with minimal false positives.

Each policy includes detailed evaluation criteria, explicit exception rules, and calibration examples that enable nuanced contextual analysis. This approach handles edge cases gracefully and adapts to natural language variations.

Flexible LLM Architecture

The platform supports multiple LLM deployment models to match organizational requirements:

Cloud LLM

- Higher accuracy for complex policy evaluation
- Suitable for standard enterprise deployments
- Faster evaluation times

Local LLM (On-Premise)

- Data never leaves the organization's infrastructure
- Required for air-gapped or highly regulated environments
- Supports industry-standard local LLMs

6. Policy Framework

Policy Categories

Whiteout AI includes 65 pre-built policies across 9 regulatory categories:

Category	Policies	Description
PHI	10	Protected health information for HIPAA compliance
PII	12	Personally identifiable information across industries
Code/IP	8	Source code, credentials, and trade secrets
Legal	6	Attorney-client privilege and litigation protection
Confidential	8	Business-sensitive internal information
Finance	4	Financial data and unreleased earnings
GDPR	7	EU data protection regulation compliance
Security	7	Infrastructure credentials and vulnerability data
Education	3	Student records (FERPA) and research data

The PHI category covers patient names, diagnoses, medical records, prescriptions, lab results, insurance identifiers, genetic data, mental health records, clinical images, and appointment details. PII policies address Social Security numbers, credit cards, addresses, phone numbers, email addresses, passport numbers, and more. Code/IP policies protect API keys, database credentials, proprietary source code, encryption keys, and infrastructure configurations.

Policy Groups

Organizations can create policy groups to apply different policies to different teams:

- **Engineering:** Focus on code/IP protection
- **Healthcare Workers:** Strict PHI enforcement
- **Finance Team:** Financial data and confidential business policies
- **Legal Department:** Attorney-client privilege and litigation protection
- **Default:** Balanced policy set for general employees

Policy groups support inheritance, allowing organization-wide defaults with group-level overrides.

Tailored Policies

Beyond the pre-built library, organizations can request tailored policies built by the Groovy Security team using the same intelligent evaluation framework. Tailored policies can address organization-specific compliance rules for proprietary projects, internal code names, or industry-specific requirements — with the same contextual understanding as the built-in policy library.

7. Interception Technology

Browser Extension

The browser extension uses a sophisticated interception architecture to capture AI interactions without disrupting user experience:

- **Interface Detection:** Automatically identifies AI chat interfaces
- **Event Interception:** Captures prompt submission before it reaches the AI service
- **Content Extraction:** Extracts the prompt text for compliance evaluation
- **Compliance Evaluation:** Sends the prompt to the backend for real-time policy check
- **Conditional Submission:** Original submission proceeds only if compliant

File uploads are intercepted across all methods — drag-and-drop, paste, and file picker — and scanned for policy compliance before reaching the AI service.

The extension includes a cross-browser compatibility layer that provides consistent behavior across Chrome, Firefox, Edge, and Safari.

Desktop Guard

Desktop Guard uses native OS-level APIs on both macOS and Windows for keyboard interception, text extraction from AI application interfaces, visual overlay for file scanning, and secure credential storage using platform-native mechanisms.

Fail-Safe Design

A critical design principle of Whiteout AI is that technical failures never block productivity. If any component experiences a temporary issue, users are never blocked from their work. The system provides clear visual feedback about compliance status — using distinct indicators for approved, blocked, and degraded states — while ensuring productivity is never impacted by technical failures.

8. Security Architecture

Authentication & Authorization

Whiteout AI uses industry-standard token-based authentication with short-lived access tokens, rotating refresh tokens, and device binding for enhanced security. All authentication flows enforce strong password policies and are protected against abuse with intelligent rate limiting.

SSO/IDP Integration

Enterprise identity providers are supported for seamless authentication:

- **SAML 2.0:** Okta, Azure AD (Entra ID), OneLogin, Ping Identity
- **OAuth 2.0:** Google Workspace, Microsoft 365
- Automatic organization detection via email domain mapping
- Just-in-time user provisioning on first SSO login

Credential Protection

All stored credentials (OAuth tokens, API keys) are encrypted at rest using industry-standard symmetric encryption with mandatory production key management. Client applications use platform-native secure storage — macOS Keychain, Windows Credential Manager, and browser-provided secure storage — ensuring credentials never persist in plaintext.

API Security

All API endpoints are protected with intelligent rate limiting, strict origin allowlists, and TLS encryption in transit. The platform enforces strong password policies and implements comprehensive input validation across all endpoints.

Audit & Logging

All security-relevant events are logged in structured format capturing timestamp, event type, user context, evaluation result, risk score, and source. Sensitive data such as prompts, tokens, and credentials are never included in log output.

Organizations can configure real-time event delivery to their SIEM platform through signed webhook integration. Supported delivery modes include real-time, batched, and filtered (high-risk events only).

9. Integration Ecosystem

External Tool Integrations

Whiteout AI connects to external productivity tools, enabling AI assistants to safely access organizational data through authenticated, audited channels:

- GitHub — Issues, pull requests, code review
- Jira — Project management and issue tracking
- Confluence — Documentation and knowledge base
- Slack — Team communication and channels
- Google Drive — Documents and file management
- And more — SharePoint, Asana, Notion, Microsoft Teams, and additional integrations

Each integration has configurable access controls determining which AI surfaces (internal AI, external tools) can access organizational data.

Developer Tool Integration

Through the Model Context Protocol (MCP), external AI development tools can securely access organizational integrations with full authentication, access control, and audit logging. This enables developers to use AI-powered coding tools while maintaining governance over data access.

Data Sovereignty

For organizations requiring full data sovereignty, Whiteout AI supports a hybrid architecture where LLM inference runs locally on the user's machine while only structured data requests are handled server-side. Prompt content never leaves the organization's infrastructure.

10. Deployment Models

Cloud Deployment

Standard deployment with Whiteout AI backend hosted in the cloud with managed database infrastructure. Desktop applications are distributed as native installers, and the browser extension is published to browser extension stores.

On-Premise Deployment

Full on-premise deployment for highly regulated industries. All components — backend, database, and LLM inference — run within the organization's infrastructure using standard container orchestration. This deployment model is suited for air-gapped environments and organizations with strict data residency requirements.

Hybrid Deployment

A hybrid model where the core API runs in the cloud while compliance evaluation happens on-premise. Prompts are evaluated locally — only metadata is sent to the cloud. This provides the benefits of cloud management while ensuring prompt content never leaves the organization's network.

11. Enterprise Use Cases

Healthcare: HIPAA Compliance

Challenge: Healthcare organizations need AI productivity while preventing PHI disclosure.

Solution: Deploy Whiteout AI with PHI-focused policies:

- 10 PHI-specific policies covering all HIPAA identifiers
- Real-time blocking of patient names, medical record numbers, and diagnoses
- Aggregate statistics and operational queries remain allowed
- Complete audit trail for compliance audits

Financial Services: Insider Trading Prevention

Challenge: Prevent accidental disclosure of material non-public information (MNPI).

Solution: Deploy with Finance and Confidential policies:

- Block unreleased earnings data
- Block M&A; details before public announcement
- Block financial forecasts and projections
- Allow public financial analysis and general queries

Technology: Source Code Protection

Challenge: Developers want AI assistance without exposing proprietary code.

Solution: Deploy with Code/IP policies:

- Block API keys, database credentials, and encryption keys
- Block proprietary algorithms and trade secrets
- Allow general coding questions and open-source discussion

Legal: Privilege Protection

Challenge: Law firms need AI for document drafting without waiving privilege.

Solution: Deploy with Legal policies:

- Block attorney-client communications
- Block case numbers and litigation strategy
- Allow general legal research queries

12. Regulatory Compliance

HIPAA (Health Insurance Portability and Accountability Act)

Whiteout AI supports HIPAA compliance through:

- Access controls ensuring only authorized personnel interact with PHI

- Complete audit trail of all AI interactions involving health data
- Transmission security preventing PHI from reaching external AI services
- Minimum necessary enforcement — only aggregate data is permitted

GDPR (General Data Protection Regulation)

Whiteout AI supports GDPR compliance through:

- Data minimization — blocking personal data from external processing
- Purpose limitation — AI interactions logged with justification
- Accountability — demonstrable compliance through audit trails
- Data subject rights support through comprehensive logging

SOX (Sarbanes-Oxley Act)

Whiteout AI supports SOX compliance through:

- Internal controls over financial data in AI interactions
- Immutable audit trail for all financial data access
- Access controls restricting financial data by role

PCI-DSS (Payment Card Industry Data Security Standard)

Whiteout AI supports PCI-DSS compliance through:

- Blocking cardholder data from reaching external AI services
- Access control measures limiting who can interact with payment data
- Complete tracking and monitoring of all data access

13. Conclusion

The enterprise adoption of generative AI is inevitable — the productivity benefits are too significant to ignore. However, organizations cannot sacrifice security, compliance, and data protection in pursuit of AI-driven productivity.

Whiteout AI provides the governance layer that enables safe AI adoption:

- **Universal Coverage:** Browser, desktop, and internal AI interfaces
- **Intelligent Compliance:** AI-powered policy evaluation with 99.5% accuracy
- **Minimal Friction:** Sub-second evaluation with fail-safe design
- **Enterprise Ready:** SSO, SIEM integration, complete audit trails

By deploying Whiteout AI, organizations can confidently enable AI tools for their workforce while maintaining the security and compliance posture their stakeholders demand.

Document Version: 2.0 | Last Updated: February 2026 | Classification: Public

Groovy Security, Inc. | groovysec.com | security@groovysec.com