

Open Source, Self-Hosted PII Redaction

The Philterd toolkit finds and removes sensitive information from text, documents, LLM prompts, and data stores. Every component runs inside your cloud, under your control, with full source code on GitHub.

Data Sovereignty

Runs inside your VPC. No data ever leaves your perimeter, reaches a third-party API, or lands in someone else's logs.

Open Source Integrity

Apache 2.0 licensed. Your engineers can read every line, audit every decision, and extend the stack on their own terms.

Purpose-Built AI

Specialized NLP models trained for PII and PHI detection. Accurate, tunable, and affordable at scale.

THE TOOLKIT

Philter	Turnkey, self-hosted PII redaction API. Drops into any pipeline that needs sensitive data removed from text.
Phileas	Core redaction library in Java, Python, .NET, and Go. Embed directly into your application with no service to deploy.
PhEye	AI and NLP model server. Purpose-trained models for PII/PHI entity recognition with domain-specific lenses.
Philter AI Proxy	Drop-in proxy that redacts PII from prompts before they reach LLM providers like OpenAI and Anthropic.
Phinder	Discovery scanner that crawls files and storage to map where sensitive information lives across your environment.
Phield	Monitoring that tracks PII flow across the organization and alerts on suspicious activity or unusual trends.
Arbiter	Human-in-the-loop review. Search, review, and override automated detection decisions with structured exemption codes.
Philter Scope	Benchmarking tool that scores redaction policies on precision and recall against gold-standard test data.
Philter Diffuse	Differential privacy for PII analytics. Preserves statistical utility without exposing individuals.
Redaction Policy Editor	Visual, no-code policy builder for non-technical users. Exports valid Philter-compatible JSON policies.

INDUSTRIES

Healthcare

Government

HIPAA Safe Harbor automation across all 18 identifiers. Healthcare NLP lens for clinical text. BAA-friendly deployment in your existing cloud environment.

Financial Services

PCI DSS scope reduction with Luhn-validated PAN detection. GLBA NPPI handling for customer data. Call-center transcript redaction for spoken card numbers.

Legal

FRBP 9037 and FRCP 5.2 automation. Human-in-the-loop review via Arbiter. Per-document audit trails for court filings and e-discovery productions.

FOIA response automation. Deploys to AWS GovCloud, Azure Government, or fully air-gapped. Open source so your security team can defend every detection path in writing.

AI and Machine Learning

LLM prompt guardrails via the AI Proxy. Training-data de-identification. RAG pipeline redaction to keep PII out of vector stores.

Education

FERPA compliance for student records. Detection of student names, IDs, dates of birth, and education-specific identifiers.

COMPLIANCE COVERAGE

- HIPAA Safe Harbor (45 CFR 164.514)
- HIPAA Expert Determination
- GDPR (EU 2016/679)
- CCPA / CPRA
- PCI DSS v4.0
- GLBA (Gramm-Leach-Bliley)
- FRBP 9037 (Bankruptcy)
- FRCP 5.2 (Federal Civil)
- FERPA
- SOX (Section 802)
- FedRAMP (NIST 800-53)
- CMMC
- ITAR

DEPLOYMENT OPTIONS

- **Cloud VPC:** AWS, GCP, Azure. Marketplace listings for one-click Philter deployment.
- **GovCloud:** AWS GovCloud, Azure Government for FedRAMP Moderate and High.
- **Air-gapped:** Self-contained Docker images. No outbound network required.
- **On-premises:** Docker or bare-metal. No SaaS dependency.

GETTING STARTED

Open Source (Free)

Run the entire toolkit yourself. Full source on GitHub. No license keys, no usage caps.

Cloud Marketplaces

Deploy Philter from the AWS, Google Cloud, or Azure marketplace. Production-ready in minutes, billed per hour through your existing cloud account.

Ready to evaluate? Visit www.philterd.ai or contact us at www.philterd.ai/consulting to schedule a 30-minute review of your data flows and compliance requirements.