Lopam DSPM

Find sensitive data, see who can reach it, and act on the exposure — across files and databases alike.

Lopam DSPM discovers structured and unstructured data stores, classifies their contents against predefined and custom policies, maps the identities and permissions that touch sensitive data, and drives a guardrail, alerting and remediation-approval workflow on top — with predefined coverage for India's DPDP Act, PCI card data and general PII out of the box. Data plane, control plane and console deploy together as one stack.

VERSION	DEPLOYMENT	DATA SOURCES	FILE COVERAGE
2.0	**On-prem or SaaS · Docker Compose**	**Filesystem + 7 database / warehouse engines**	**72 recognized file types**

01 Architecture

Three components against one shared JSON contract. The engine runs near the data; the control plane holds every model, policy and workflow; the console is where analysts work.

DISCOVERY ENGINE

dspm-engine

Filesystem walker with no depth limit. File-type detection, text extraction, content classification, staleness scoring and ACL capture. Runs as a CLI: **scan** or real-time **watch**.

`Go 1.21 · standard library only · shells to tesseract / pdftoppm / yara when present`

CONTROL PLANE

Django backend

Every entity and the full REST API under **/api/v1/dspm/**. RBAC with TOTP MFA, structured-data scanning, the guardrail engine, SIEM dispatch, and Celery-scheduled reports and staleness sweeps.

`Django 5 + DRF · Celery + Redis · PostgreSQL`

CONSOLE

React SPA

Dashboard, Data Stores, Scans with live progress, Findings, Policies, Access Intelligence graph, Guardrails, Remediation, Alerts, Reports, Integrations and Admin/RBAC.

`React 18 + TypeScript · Vite · session-cookie auth`

Data store → Engine / structured connector → Findings + permissions → Guardrails → Alerts / remediation → SIEM · STIX/TAXII

02 Discovery & classification

PREDEFINED POLICIES	India **Aadhaar** (12-digit), **PAN**, **PCI** card number with real Luhn validation, email address, India / US phone, IP address.
CUSTOM PATTERNS	`regex` `keyword` `dictionary` `boolean AND/OR/NOT` `proximity` `YARA rule`
CONTEXT SIGNALS	Label-proximity confidence boosting and, for structured data, column-name matching — a column literally named `aadhaar` or `card_number` lifts confidence on its own.
AI/ML PASS	A lightweight scoring model reads keyword density and text entropy to catch sensitive content that pattern matching alone misses, and reports its own findings under a distinct **ai_ml** source so analysts can tell a model call from a policy hit.
OCR	Images and image-only PDF pages via tesseract + pdftoppm when present on the host; otherwise the finding records **ocr_unavailable** rather than failing.
METADATA & FILTERS	Discover by creation / last-modified date, file size, file type, file name, and document header / footer text.
MULTI-TYPE FINDINGS	A single file reports every sensitive type it contains — PCI and PII and HIPAA together, not just the first match.

03 Sensitivity model

Every finding resolves to one level. The console’s dashboard, filters and permission graph are all keyed to this scale.

critical	high	medium	low	none
Regulated identifiers in combination — e.g. Aadhaar + card number	Strong regulated identifier present (Aadhaar, PAN)	Contact-grade PII — email, phone in volume	Weak or isolated signals	Scanned, nothing sensitive matched

04 Data source coverage

Unstructured

Filesystem stores, unbounded nesting. 72 reference file types across documents, spreadsheets, presentations, email, archives, images and code.

`native extraction` `legacy-format extraction` `OCR on images & scans` `archive traversal`

Structured

Connector-based schema / table / column enumeration with row sampling and the same policy evaluation as the filesystem path.

`PostgreSQL` `MySQL` `MSSQL` `Oracle` `Snowflake` `BigQuery` `Redshift`

SCAN MODES	**Full**, **incremental** (real-time via filesystem watch — not scheduled), and **on-demand scoped** scans that run alongside an environment-wide scan without disturbing it.
PROGRESS	Live files-discovered count, files-scanned count, percent complete and flag count, streamed to the console.
STALENESS & RETENTION	Configurable age / last-access thresholds, plus retention policies such as “HR documents older than 3 years containing PII.”

05 Access intelligence

Permission graph

Interactive force-directed and hierarchical views of which identities and groups can reach which data objects, with filters for protected or unique folders.

Exposure analysis

- Every user with access to a given sensitive data type, and their effective rights
- Stale identities that still hold access
- Unused / stale permissions on sensitive data
- Third-party application access and the identities behind it

- Full-control / create / read / write / delete / list breakdown per store

Permission changes are re-mapped on the next scan of the store, and classification context is carried into permission reports.

06 Guardrails, alerting & remediation

GUARDRAIL ENGINE	Admin-console policy layer that evaluates findings in real time and fires Alerts with actions of **alert**, **alert & auto-remediate**, or approval-gated remediation.
REMEDIATION ACTIONS	Remove org-wide permissions, remove stale permissions, revoke a specific grant, or move / copy stale data to a quarantine location — natively, with no third-party tool in the path.
APPROVAL WORKFLOW	Every remediation action routes to the accountable business owner for approve / reject before it runs, with the full decision recorded for audit.
ALERT POLICIES	Guardrails evaluate every scan finding and permission event as it happens, not on a schedule. Teams commonly configure them for stale-data access, atypical-folder access, unusual external exposure and ransomware-note patterns.
AUDIT LOG	Data activities — creation, modification, deletion, sharing — and all administrator activity are recorded for investigation.

07 Integrations & API

SIEM / log manager

`Syslog` `CEF` `Webhook (JSON)` `STIX 2.1` `TAXII 2.1 (read)`

Alerts dispatch as CEF over Syslog and as JSON webhooks, and surface as STIX Indicator objects on a TAXII collection endpoint.

Delivery

Alerts push the moment a guardrail fires — no polling. Configure any number of Syslog and webhook targets per tenant from the console.

REST API /api/v1/dspm/

```
# core resources
/api/v1/dspm/data-stores/   scan-jobs/   findings/   permissions/   identities/   third-party-apps/
/api/v1/dspm/guardrail-policies/   alerts/   remediation-actions/   reports/   audit-log/
# dashboards, auth, integrations
/api/v1/dspm/dashboard/summary/   auth/login/   auth/mfa/setup/   auth/mfa/verify/
/api/v1/dspm/siem-config/   siem-targets/   taxii2/dspm/collections/
```

08 Security, governance & reporting

Access control

- Native TOTP multi-factor authentication for administrators
- Comprehensive role-based access control
- Dedicated role gate for viewing actual file contents
- Detailed administrator-activity auditing

Reporting

- Discovery-task tracking and multi-type findings per file
- PII created / modified / deleted on a date or range
- Attribute grouping and add / remove customisation
- CSV, XLSX and PDF export
- Scheduled delivery by email on saved filters

09 Deployment

ON-PREMISES / PRIVATE CLOUD	The full stack — PostgreSQL, Redis, the API, background workers and the console — deploys as one Docker Compose stack behind your own perimeter.
MANAGED SAAS	The same stack, operated for you, for teams that would rather not run the infrastructure themselves.
BRING YOUR OWN DATA	Point a data store at a filesystem path or a database connection and the first scan starts immediately — no schema migration or agent install on the source system.

Lopam DSPM — data security posture management for structured and unstructured data.

Ask your account team for a walkthrough on your own data stores.