

Data Loss Prevention (DLP) and data discovery

Know where your sensitive data is, and who can reach it

Finds Aadhaar numbers, PAN, bank accounts, card numbers, keys and your own confidential records where they should not be — on servers, in OneDrive, SharePoint and Google Drive, and in mail that leaves — and takes access away with a second person's approval.



TotalHeal
DataShield

Contents

1	Executive summary	6	Security and data handling
2	The challenge	7	Where it fits
3	Architecture	8	Deployment and system requirements
4	How it works	9	Technical specifications
5	Capabilities in detail	10	Getting started

1 Executive summary

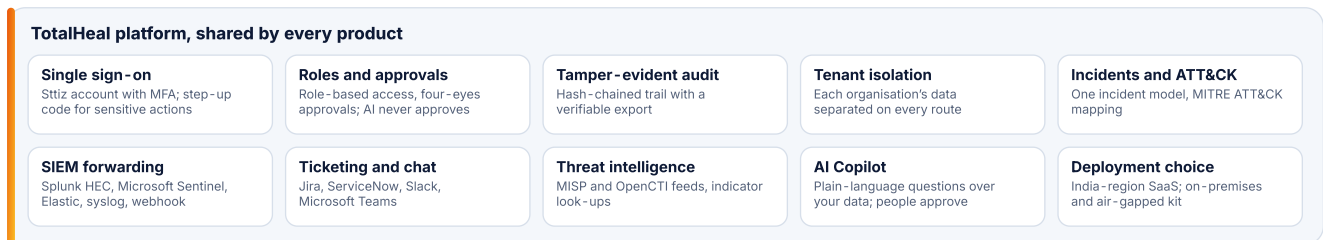
DataShield inspects servers, OneDrive, SharePoint, Google Drive, sent mail and anything sent to its scan API with identifiers validated by checksum, your own records matched exactly and confidential documents recognised by fingerprint; it shows who can reach each finding and takes access away when a second person approves. Sensitive values never leave the inspection unmasked.

2 The challenge

Sensitive data spreads: a patient export left on an application server, a .env with production keys every user can read, a vendor list with PAN and bank accounts mailed to an outside firm, a results sheet shared with “anyone with the link”. Under the DPDP Act 2023, the Aadhaar Act and PCI DSS the organisation must know where this data is and who can reach it — and pattern-matching tools drown teams in invoice numbers that look like account numbers.

The TotalHeal approach. DataShield inspects files on Linux servers, in Microsoft 365 and Google Workspace, in sent mail and anything sent to its scan API with an engine of its own: identifiers validated by their checksums, your own records matched exactly, confidential documents recognised by fingerprint. It shows where each finding is and who can reach it, and removes public links or file access when a second person approves.

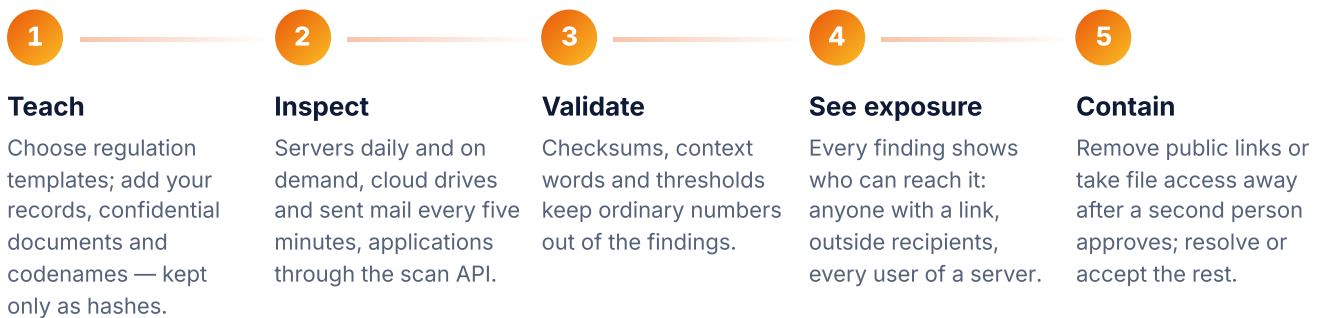
3 Architecture



TotalHeal component
Your environment
TotalHeal cloud or edge
Third party
 Data
 Control, optional or on approval

Servers inspect files locally and send only masked findings; cloud files and sent mail are read through the providers' APIs and forgotten after inspection; access changes run only after a second person approves.

4 How it works



5 Capabilities in detail



25 detectors, validated

Aadhaar (Verhoeff check digit), PAN (holder type), GSTIN (check character), bank account numbers next to account words, IFSC, UPI IDs (known bank and app handles), Indian mobile numbers, passport, voter ID (EPIC) and driving-licence numbers, payment cards (Luhn and real network prefixes: Visa, Mastercard, RuPay, Amex, Discover, Diners, JCB), IBAN (mod-97), SWIFT, US SSN, e-mail addresses — each with context words that raise its confidence; repeated and sequential test numbers are ignored.



Exact data match

Upload a CSV of your customers, patients, employees or accounts: DataShield keeps only salted hashes and finds files holding two or more fields of the same record, however the numbers are written.



Keyword dictionaries

Codenames, project names and labels such as “Strictly confidential”, matched as whole words in any case.



Policies for regulation

Ready templates for the Aadhaar Act, PCI DSS, the DPDP Act 2023 (bulk personal data), Indian financial identifiers, credentials and GDPR / GLBA, and your own policies: data types with minimum counts and confidence, your records, documents and words, per channel — edited in the console, never as JSON.



OneDrive, SharePoint and Google Drive

Changed files in every user's drive (and every SharePoint site library) are inspected every five minutes, with how each is shared: a link anyone can open, people outside your domains, the whole organisation or nobody. Google Docs, Sheets and Slides are inspected through their text.



Secrets and keys

Private keys (RSA, EC, OpenSSH, PGP), AWS access keys, GitHub, Slack and Google API tokens, Stripe and Razorpay live secret keys, JSON Web Tokens, database and broker connection strings with passwords, Azure storage keys and passwords in configuration files (placeholders ignored).



Document fingerprints

Register a board paper, contract or design: files that copy a share of its text are found, even with other text around it. Only hashes of the document are kept.



Text from real files

Word, Excel and PowerPoint (Office Open XML), OpenDocument, PDF text, CSV, JSON, XML, HTML, e-mail with attachments, and zip archives nested up to three levels; legacy Office files by their text; password-protected files reported, never silently skipped; archive bombs refused.



Servers: data at rest

The TotalHeal server agent inspects the folders you choose on every Linux server every day and on demand, reads only new or changed files, never follows links and never reads system trees, and reports who owns each file and whether every user of the server can read it.



Mail that leaves

Every sent message and its attachments in Microsoft 365 and Gmail is inspected; recipients outside your domains make the finding exposed.



Scan API

POST any file or text to the scan API from your applications, upload paths or CI and get the verdict, the masked findings and the policies broken; record the result as a finding if you choose.



Masked evidence

A full sensitive value never leaves the inspection: findings keep the count, the confidence and up to three masked samples (..... 9012), the file and the part of an archive or attachment it was in.



Take access away, with approval

On a server, group and other users lose access to a flagged file and its old permissions are journaled so it can be put back; in Microsoft 365 or Google Drive, public links and outside shares are removed while shares inside the organisation stay. Every change needs a second person — never the requester, never a service account.



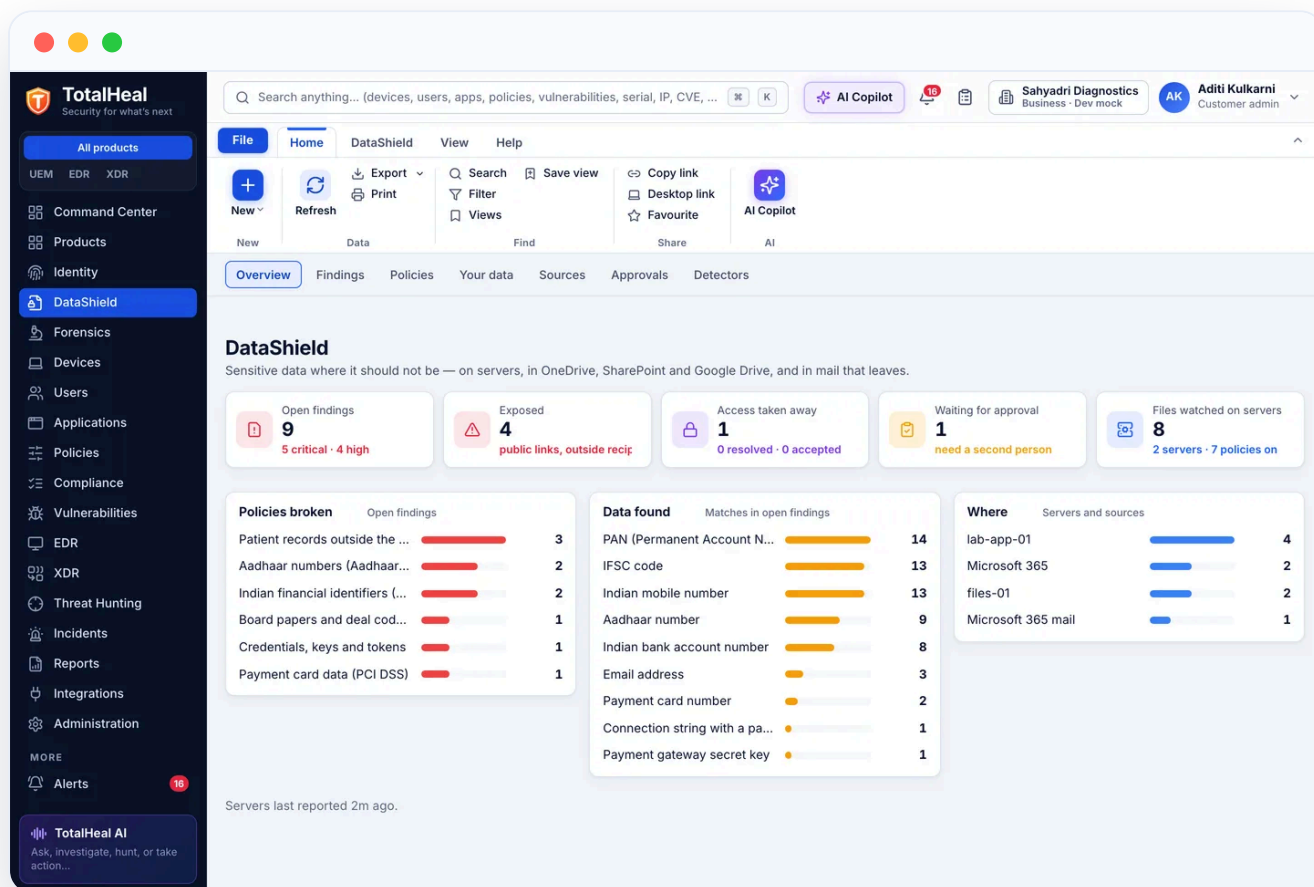
Findings that stay current

Resolve, accept the risk with an expiry, or reopen; files that are deleted or cleaned are resolved by the next scan, and a finding that comes back reopens.



Alerts in the same incident queue

A new finding that breaks a policy becomes an alert (credentials in files: MITRE ATT&CK T1552.001) correlated with endpoint, identity and cloud alerts and sent on to your SIEM, ticketing and playbooks.



The DataShield overview in the TotalHeal console (demo organisation). Real screenshot, demo data.

6 Security and data handling

Values stay masked

Only counts and masked samples leave the inspection; records and documents are kept as hashes.

Every server change can be undone

The old permissions of a restricted file are journaled and put back on request.

Two people for every change

Taking access away or removing a link needs a second person — never the requester, never a service account.

Read-only until you allow it

Servers change nothing unless the host owner allowed access changes.

7 Where it fits

DPDP Act readiness

Find personal data of many people outside the systems meant to hold it, with evidence of where and who can reach it.

Aadhaar and card data

Prove Aadhaar numbers and card numbers are not lying in exports, uploads and shared folders.

Leaked keys

Find production keys and passwords in files every user can read, and take access away the same day.

8 Deployment and system requirements



Cloud (SaaS)

Available

Single-region service.



On-premises

Available

Delivered with our team from the on-premises kit.



Air-gapped

Available

Offline kit, offline licence file and offline update bundles.

Protected platforms



Ubuntu
20.04, 22.04,
24.04



Debian
11, 12



**Red Hat
Enterprise
Linux**
8, 9



**Rocky
Linux**
8, 9



AlmaLinux
8, 9



**Microsoft
365**



**Google
Workspace**



Gmail

On-premises server

Operating system

RHEL, Rocky Linux or AlmaLinux 8 / 9; Ubuntu 20.04, 22.04 or 24.04; Debian 11 / 12

Processor

4 vCPU (tested configuration)

Memory

8 GB recommended (6 GB minimum)

Disk

10 GB free for software, releases and backups, plus data

Sign-in

Active Directory / LDAP, SAML 2.0 or OpenID Connect

Software

Network and access

- The TotalHeal server agent on each Linux server, with `datashield.enabled` and the folders to inspect
- `datashield.allow_restrict` on servers where approved access changes may run
- For cloud files and mail: an administrator who grants API access to Microsoft 365 or Google Workspace

Alerts and tickets to

splunk> **Splunk**

**Microsoft
Sentinel**

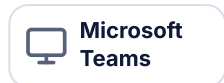
Elastic

**Syslog /
CEF**

Jira

ServiceNow

Nothing to pre-install: no container runtime, database, web server or Java



9 Technical specifications

Features	Included	
25 detectors, validated	✔ Included	
Secrets and keys	✔ Included	
Exact data match	✔ Included	
Document fingerprints	✔ Included	
Keyword dictionaries	✔ Included	
Text from real files	✔ Included	
Policies for regulation	✔ Included	
Servers: data at rest	✔ Included	
OneDrive, SharePoint and Google Drive	✔ Included	
Mail that leaves	✔ Included	
Scan API	✔ Included	
Masked evidence	✔ Included	
Take access away, with approval	✔ Included	
Findings that stay current	✔ Included	
Alerts in the same incident queue	✔ Included	
Availability by edition	Business	Enterprise
Detectors, exact data match, fingerprints, dictionaries, policies, servers, OneDrive / SharePoint / Google Drive, sent mail, scan API, approved access changes, alerts	✔ Included	✔ Included
Platforms and coverage		
Linux servers (Ubuntu, Debian, RHEL, Rocky, AlmaLinux)	TotalHeal server agent with datashield.enabled and the folders to inspect	

Microsoft 365	OneDrive, SharePoint and sent mail through Microsoft Graph (Files.ReadWrite.All, Sites.Read.All, Mail.Read, User.Read.All), or keyless with a federated credential
Google Workspace	Drive and Gmail sent mail through a service account with domain - wide delegation
Applications and pipelines	The scan API (any file or text up to 25 MB)
Console	TotalHeal web console
Deployment options	
SaaS (India region)	Available Single - region service.
On - premises control plane	Available Delivered with our team from the on - premises kit.
Air - gapped	Available Offline kit, offline licence file and offline update bundles.
Licensing	
Priced per	Server per year
Counted	Servers reporting to DataShield; Microsoft 365 / Google Workspace connections included
Trial	30 days, up to 5 servers
Editions	DataShield is sold on its own, outside the editions.
Government profile	Government is a profile, not a separate tier or price list: the Enterprise tier with Government defaults switched on — data residency in India, four - eyes approval of policy changes, a distinct approver for every approval - level action, retention of at least 180 days, and the CERT - In and DPDP reporting clocks.
Procurement dossier	Compliance matrix, bill of materials, architecture and hardening, SLA and OEM authorisation letter, on request
On - premises server	
Operating system	RHEL, Rocky Linux or AlmaLinux 8 / 9; Ubuntu 20.04, 22.04 or 24.04; Debian 11 / 12
Processor	4 vCPU (tested configuration)
Memory	8 GB recommended (6 GB minimum)
Disk	10 GB free for software, releases and backups, plus data
Sign - in	Active Directory / LDAP, SAML 2.0 or OpenID Connect
Software	Nothing to pre - install: no container runtime, database, web server or Java
Requirements	
Requirement 1	The TotalHeal server agent on each Linux server, with datashield.enabled and the folders to inspect

Requirement 2	datashield.allow_restrict on servers where approved access changes may run
Requirement 3	For cloud files and mail: an administrator who grants API access to Microsoft 365 or Google Workspace

10 Getting started

- The TotalHeal server agent on each Linux server, with datashield.enabled and the folders to inspect
- datashield.allow_restrict on servers where approved access changes may run
- For cloud files and mail: an administrator who grants API access to Microsoft 365 or Google Workspace

Start a 30 days, up to 5 servers trial in the TotalHeal console with a Sttiz account. For a quote, an on-premises installation or the procurement dossier (compliance matrix, bill of materials, architecture and hardening, SLA, OEM authorisation letter), write to business@totalhealsecurity.com.

About TotalHeal

TotalHeal is a security platform from Freaquer Corporation Pvt Ltd: endpoint, network, application, cloud, security-operations and backup products, each sold on its own and run from one console.

Datasheet

totalhealsecurity.com/datasheet/datashield

No third-party certification is claimed. Figures are from our own tests where stated, with their dates. Specifications can change; this whitepaper is dated 27 October 2026.