

AI THAT KNOWS YOUR BUSINESS

Nasuni AI Activate Solution Brief

Give AI apps governed, permission-aware access to enterprise file data.

The data AI needs most is the data it can't reach

More than 90% of enterprise data lives in files: contracts, specifications, reports, project records, and correspondence. That is the institutional knowledge your teams depend on, and the knowledge your AI cannot reach. Most pilots demo well against sample data, then stall before production.

The obstacle is rarely the model. It is the path between the model and the file estate, where content is large, permission-controlled, and spread across sites and clouds. Nasuni AI Activate closes that gap.

90%+ of organizational data is unstructured	94% of enterprises struggle to manage unstructured data	97% are piloting AI	18% have reached enterprise scale with it
---	---	-------------------------------	---

Sources: McKinsey, *Intelligence at Scale* (90%+). Nasuni, *State of Enterprise File Data 2026*, n=1,000 (94%, 97%, 18%).

/ What AI Activate is

Nasuni AI Activate is a hosted service enabling AI apps with a governed, live access to enterprise file data, enforcing your existing permissions before any query reaches a file. No copies into the AI app, and no new pipelines to build.

It consists of two components, plus a read-only path to your data:

- **Indexing service.** Reads file content and metadata from the Nasuni volumes you scope and builds a searchable index.
- **MCP Server.** Identifies the authenticated user, passes that identity to the index, and returns only the content that user is permitted to see.

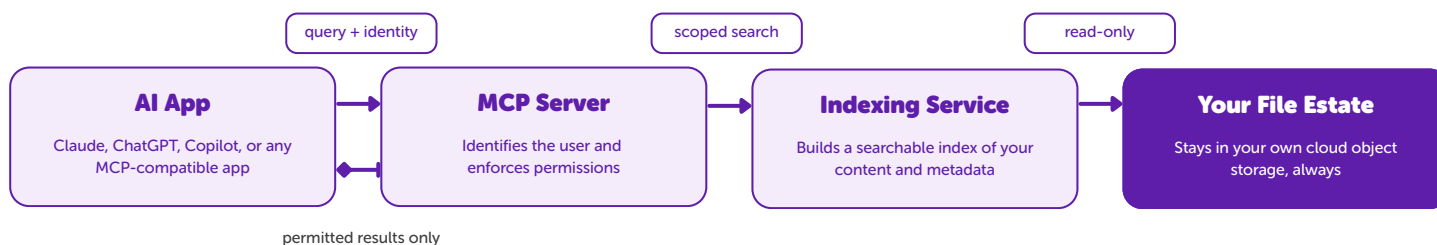


Diagram 1 – Architecture. Two hosted components (MCP Server, Indexing Service), plus a read-only path to your data: No copies, no write-back.

/ What AI Activate is not

- **It does not move or copy your files.** Your original files stay in your own cloud object storage.
- **It does not write to your data.** AI Activate reads only, and never alters or deletes files.
- **It does not train AI models on your content.** Nasuni never uses your content to train or fine-tune any AI model.

What teams do with it

Here are three ways teams use governed AI access to unstructured data. Each example reasons across content, not just keywords.

/ Find

“Find vendor contracts with an auto-renewal clause expiring in the next 90 days, across every regional office.”

Clause-level retrieval across the whole estate, not a filename match in one folder. Renewal exposure surfaces while there is still time to act, and the answer covers every site rather than the ones someone remembered to check.

/ Analyze

“Compare this specification against every incident report filed since it was last updated, and flag what looks related.”

Reasoning across many documents at once, cross-referencing the specification against every incident record tied to it. The pattern that no single document contains becomes visible.

/ Build

“Draft a customer-ready incident summary from the maintenance logs, inspection notes, and warranty claims tied to this asset.”

Synthesis from every relevant source into a review-ready draft. The gathering work that normally comes first is already done, so your team starts from your own records rather than a blank page.

What governed access delivers

Seven properties turn a file estate into data that knows your business. Enterprise AI needs all seven at once, and the weakest one sets the ceiling on every answer it gives you. The seven spell PSYCHIC: Proprietary, Secure, Yours, Current, Hybrid, Indexed, and Continuous.

	Property	What AI Activate delivers	Without it
P	Proprietary	Data only your organization has: contracts, specifications, reports, and correspondence.	Generic answers any competitor’s AI can produce.
S	Secure	Permissions enforced at retrieval, before any content is returned.	A compliance project where you wanted a productivity gain.
Y	Yours	No second copy of your data in any third-party AI or analytics platform.	A copy of your data in someone else’s platform, with its own permissions to rebuild and its own exposure surface.
C	Current	The live file and its current permissions, not last week’s export.	Confident answers built on a stale snapshot.
H	Hybrid	One namespace across every site, cloud, and edge location.	AI that works at headquarters and goes blind in the field.
I	Indexed	The right file in seconds across the entire estate.	The answer exists, and the model guesses instead of finding it.
C	Continuous	Immutable version history is the foundation AI Activate builds on. AI reasons over your files’ current state today.	No history of document evolution, and the last save is all that exists.

Nasuni AI Activate delivers all seven together, on one platform: built in, not bolted on.

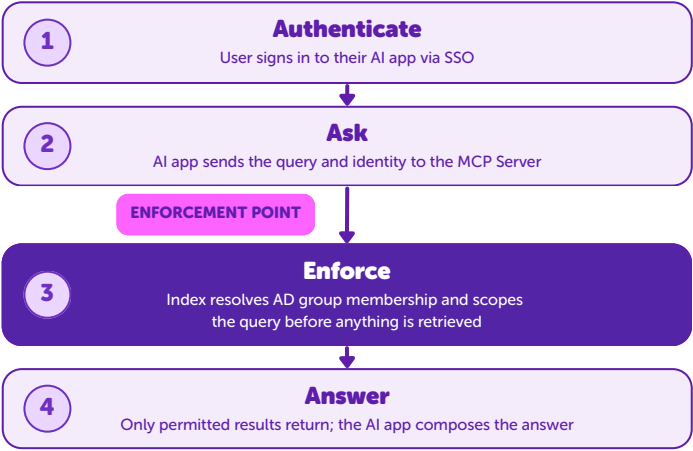
Governed AI access to unstructured data — in weeks, not quarters.

How it works

A single question travels four steps:

- 1 The user authenticates to their AI app through SSO.
- 2 The user asks a question. The AI app builds an MCP tool call and passes the user's identity to the MCP Server.
- 3 The index resolves the user's Active Directory group membership in real time and, most importantly, scopes the query to the content that identity is authorized to access, before anything is retrieved.
- 4 Only permitted results return to the AI app, which composes the answer.

Diagram 2 – The query lifecycle. Enforcement sits before retrieval, not after.



The one sentence to remember. Through the MCP Server, a user reaches only the files their own file-level permissions already allow. It follows from filtering at the data layer, before results return.

/ Connecting an AI app

MCP is to AI apps what USB-C is to devices: one standard port rather than a custom cable for every pairing. Any MCP-compatible app connects the same way, including Claude, ChatGPT, Microsoft Copilot, Google Gemini, and custom MCP solutions, with no per-app integration to build.

Configuration item	What it is	Where it is set
MCP Server URL	The endpoint the AI app connects to	Provided in the Nasuni Portal after onboarding
Authentication	How the app proves user identity	SSO token from your identity provider, the same directory used for the Portal
Scope	Which volumes and paths the server exposes	Set per volume in the Portal, using include and exclude path rules
Tool registration	Registering the Nasuni MCP Server with the AI app	In the AI app's own admin interface

/ What gets indexed

Coverage	What is indexed	Applies to
Content and metadata	Full-text content, plus name, path, size, type, modification date, and permissions	Microsoft Office, OpenDocument, Apple iWork, searchable PDF, HTML, XML, email (EML/MSG), EPUB, RTF, TXT, Markdown, CSV/TSV, and common source code
Metadata only	Name, path, size, type, modification date, and permissions	Every other file, including CAD, BIM, Revit, images, audio, video, and scanned or image-only PDF without OCR

Files indexed by metadata only stay fully addressable. Scanned and image-only PDFs carry no text layer, so identify those volumes early when you scope a proof of concept.

/ How the index stays current

After a one-time initial crawl of a newly scoped volume, the indexing service processes only what changes. Near-real-time delta processing picks up files as they are created, modified, and deleted, along with changes to a file's ACLs, without a full re-scan. Group memberships resolve at query time, so a group change takes effect on the next query, while an edit to a file's own ACL takes effect within your own snapshot interval.

Governed access and the control model

AI Activate applies the permission model built into the Nasuni platform, rather than a separate system for your team to configure and certify.

Only 31% of organizations centrally enforce what data AI agents can access.

Nasuni, State of Enterprise File Data 2026, n=1,000

- **Indexing service.** Reads file content and metadata from the Nasuni volumes you scope and builds a searchable index.
- **MCP Server.** Identifies the authenticated user, passes that identity to the index, and returns only the content that user is permitted to see.
- **Pre-query enforcement.** Permissions apply at retrieval, before any content is returned. AI Activate never retrieves off-limits content, so AI cannot see what the user cannot see.
- **Identity-based access.** Access follows the user’s own Active Directory identity through SSO, with no shared or over-permissioned service accounts.
- **Files never leave your object storage.** Nasuni does not copy file binaries out of your cloud storage. The indexing service reads content to build a derived index of extracted text, metadata, and ACL state, which lives in Nasuni-managed infrastructure, encrypted under Nasuni-managed keys.
- **Independently assessed.** SOC 2 Type II applies to the AI Activate service. For the current full certification list, see trust.nasuni.com.

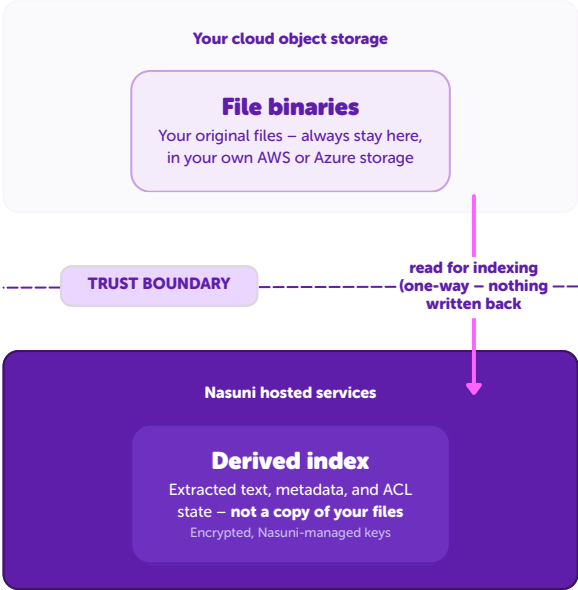


Diagram 3 – The trust boundary. File binaries never cross it. Only a derived index does.

Your files stay in your storage. What Nasuni holds is a derived index made up of extracted text, metadata, and ACL state, rather than a copy of your files.

Data type	Stays in your storage	Flows to Nasuni	Held by Nasuni
File content	Always	Read for text extraction only	Nothing retained
Extracted text and metadata	—	To the indexing service	In the index, under Nasuni-managed keys
Permissions and ACLs	In UniFS	Read at index time and query time	ACL state in the index, re-evaluated per query
Identity and SSO token	In your identity provider	Per query	Not retained
Query content and results	—	Per query	Not retained

/ Aligning AI reach with intent

Path rules bring AI-reachable content into line with what your organization intends without changing any permissions. Include and exclude path rules, set per volume in the Nasuni Portal, limit which folders the index exposes. Path rules take effect immediately and reverse cleanly — no permission changes required. Path rules are the fastest control to demonstrate in a proof of concept, and they are usually the security team’s first ask.

Why it's structurally different

/ Index-backed, not file-by-file

	Connecting an AI app straight to a file system	AI Activate
How retrieval works	Opens and reads files one at a time along a path	Answers from a purpose-built index that reflects your files' current content and metadata, without re-reading the original file at query time
As file count grows	Slows down, because the work scales with the number of files opened	Stays fast, because the work scales with the index rather than the estate
Context quality	Returns whole files, filling the model's context with material the question did not need	Returns scoped, metadata-enriched results, so the model reasons over relevant content
Permission enforcement	Depends on what the connecting account can read	Evaluated per user, per query, before retrieval

/ Pre-query enforcement, not post-query filtering

	Connecting an AI app straight to a file system	AI Activate
How it works	The index returns all matches, then filters unauthorized results afterward	The query is scoped to authorized content, so unauthorized content is never retrieved
What the model sees	All matches, including unauthorized files	Only permitted content
Compliance implication	Exposure already occurred, because the model processed restricted content even if the user never saw it	No exposure, because the model's context never contains inaccessible content
When ACLs change	Depends on how quickly the filter layer updates	Reflected at the next index pass, with no stale permission cache

What building it yourself would actually take

A capable engineering team could assemble something in this shape. The real question is what it costs to build and, more importantly, to keep running.

Role	What they build	Loaded cost/yr	Months	Potential Cost
Senior AI and ML engineer	MCP server architecture, retrieval, and reranking logic	\$250K	8	\$167K
Senior platform engineer	Indexing pipeline, search cluster, and file format handling	\$200K	8	\$133K
Security and identity engineer	Active Directory and Kerberos integration, ACL enforcement, and security controls	\$220K	6	\$110K
DevOps and site reliability engineer	Cloud infrastructure, CI/CD, monitoring, and scaling	\$190K	5	\$79K
Technical program	Architecture decisions, stakeholder management, and coordination	\$230K	8	\$77K
Total				\$566K

Representative enterprise: 6,000 employees, 14 million files, 12 offices, and US-based senior engineers. All assumptions may vary.

/ The ongoing commitment to DIY

Building it is a project. Running it is a permanent operational line item.

Category	What it covers	Potential annual cost
Engineering time	Ongoing platform ownership: index maintenance, file format updates, identity model changes, and LLM API compatibility as tool specs evolve. Minimum 0.25 FTE at steady state, 0.5 FTE for active environments.	\$60K–\$130K/yr
Cloud infrastructure	Search cluster hosting, indexing compute, and MCP server operations. Scales with file count and query volume, based on a representative 10–15 million file estate at standard cloud compute rates.	\$24K–\$48K/yr
Tooling and incident response	Monitoring, alerting, on-call coverage, and periodic security review. Excludes unplanned remediation if a permission-model issue surfaces post-deployment.	\$10K–\$20K/yr
Total	The annual cost of a system with no new capabilities delivered, purely keeping it current, secure, and operational. Variance reflects team location, file estate size, and query volume.	\$94K–\$198K/yr

What will drive the estimate up? Multi-region file estates, regulated industries, advanced compliance requirements, multiple AI app integrations, complex Active Directory environments, and third-party penetration test remediation.

That is the real cost of building it yourself: not just the project, but the permanent team, the on-call rotation, and the accountability every time a permission check fails. With AI Activate, your team can focus on the work they do best. It is enforcing permissions correctly on every query, and keeping a fresh index running as the estate changes, indefinitely.

Supported environment

Deployment	Fully hosted SaaS service on the Nasuni file data platform
Identity	SSO to the Nasuni Portal through Microsoft Entra ID or Okta
File authentication	Active Directory-authenticated volumes with NTFS permissions
Object storage	AWS or Azure-backed Nasuni volumes
Edge appliance	Nasuni Edge Appliance 10.4 or newer
AI app compatibility	Microsoft Copilot (including Copilot Studio), Anthropic Claude, OpenAI ChatGPT, Google Gemini, and any MCP-compatible AI app
Access control	Pre-query enforcement of Active Directory and NTFS permissions on every request
Scope	Volume selection, plus include and exclude path rules within a volume
Index freshness	Near-real-time delta processing
Management	Configured and scoped in the Nasuni Portal
Compliance	SOC 2 Type II applies to the AI Activate service. Full certifications: trust.nasuni.com

Nasuni evaluates fit per volume. A mixed estate that holds both eligible and ineligible volumes is normal, and it does not hold up a proof of concept on the volumes that qualify.

From first conversation to proof of concept



Diagram 4 – The path to a validated proof of concept.

Six questions to answer before your demo

Make the most of your AI Activate demo and think about the answers to these questions, as settling them first helps turn a product walkthrough into a scoped evaluation of your own estate.

- 1 Which one or two volumes hold the content a real use case depends on?
- 2 Do those volumes use Active Directory authentication with NTFS permissions?
- 3 Which AI app(s) do your teams already use, and who administers it?
- 4 Are there folders inside those volumes that AI should not reach, even where permissions would allow it?
- 5 Do broad ALLOW groups, such as Everyone or Domain Users, apply to content you plan to scope in?
- 6 What would your security team need to see in order to sign off?

Next steps

See it live

Ask your Nasuni team for a technical demo or a discussion with an AI Activate expert.

Talk to your account team

Scope a proof of concept and bring your security stakeholder in early.



Nasuni is a leading unstructured data platform for enterprises where file data is mission-critical for both people and AI. We power the operational file layer where work happens to help organizations manage, protect, and activate data so teams can work smarter, reduce costs, and operate securely without limits. Built on a patented architecture that fuses cloud object storage with enterprise file services — including permissions, versioning, and a global namespace — Nasuni delivers high-performance file access, global data availability, and a scalable, governed, AI-ready single source of truth across every major cloud. Trusted by more than 1,300+ enterprises globally, Nasuni helps organizations modernize file infrastructure, strengthen data security, and support AI-driven operations. Learn more at nasuni.com/product/ai-activate