CORNERSTONE CYBER

SAMPLE · REDACTED

AI READINESS ASSESSMENT

Make your Microsoft data safe for AI.

A worked example of the assessment output. It covers making SharePoint, OneDrive and the Microsoft 365 tenant safe before enabling any AI over corporate data: Microsoft 365 Copilot, ChatGPT Enterprise, Claude, Gemini and connected AI tools. Client and data removed.

PREPARED FOR
A global enterprise (anonymised)

ENGAGEMENT
AI readiness · Stage 1

ESTATE
Three Microsoft 365 tenants

DOCUMENT
High-Level Design · sample extract

This is a redacted sample of a real deliverable. The client, third parties and named systems have been removed and every figure has been changed so the source cannot be identified. Detailed build guidance has been held back. It shows the shape and quality of the output, not a runnable plan. Illustrative only.

Contents

The full deliverable runs to roughly 67 pages. This sample reproduces the executive and findings sections in full and abridges the design and build sections. Sections marked **abridged** are shortened or redacted here.

01 Executive summary

The organisation runs three Microsoft 365 tenants across its home region and two international regions. A data discovery exercise combining Microsoft administrative exports and data security posture management found that all three tenants carry material data governance risk, with several findings severe enough to block the planned rollout of AI over corporate data.

The risk is not any single AI product. It is enabling AI of any kind, Copilot, ChatGPT Enterprise, Claude or a connector that syncs SharePoint and OneDrive, over content never governed for AI-scale search. Copilot grounds on the tenant and inherits each user's permissions. Third-party AI reaches the same content through connectors. Both surface whatever a user can already reach. Across roughly 46,000 indexable locations, access has grown without governance, and one of seven gates is met.

~46k	~16.8k	~5.9M	1 / 7
INDEXABLE LOCATIONS IN AI SCOPE	UNAUTHENTICATED SHARING LINKS	RECORDS AT RISK TODAY	READINESS GATES MET

Rating definitions

Every finding is rated two ways. Risk level is the severity against the organisation's data and operational risk. Deployment impact is the effect on enabling AI. A finding can be high risk without blocking AI, and can block AI without being the highest severity. The two are reported independently so remediation and release decisions stay separate.

RISK LEVEL	MEANING	AI IMPACT	MEANING
CRITICAL	Immediate exposure of sensitive data or privileged access. Fix inside 14 days.	BLOCKED	AI cannot be recommended until fixed. The gate cannot clear.
HIGH	Material exposure of data or configuration. Fix inside 30 to 60 days.	CONDITIONAL	May proceed only with compensating controls and a fix commitment.
MODERATE	Localised gap contained to a specific scope. Fix this quarter.	WATCH	Monitor during rollout. Does not prevent it.
LOW	Limited exposure. Handle in standard governance cycles.	NO IMPACT	Does not affect the AI decision.

Executive findings summary

FINDING AREA	CURRENT STATE	RISK	AI
Indexable content	~18,500 SharePoint sites and ~27,500 OneDrive accounts. No ownership or classification model across either.	HIGH	CONDITIONAL
Orphaned OneDrive	~13,200 of ~24,000 drives in the largest tenant have no owner. Departed-staff content stays searchable.	CRITICAL	BLOCKED
Unauthenticated links	~16,800 Anyone links across ~290 sites. No sign-in required.	CRITICAL	BLOCKED
Data residency	~40 findings where home-region data is stored outside the region.	CRITICAL	BLOCKED
Privileged app	A third-party management app holds tenant-wide full control across all tenants.	CRITICAL	CONDITIONAL
Sensitivity labels	Around 3 percent site coverage. Purview cannot constrain what AI grounds on.	HIGH	CONDITIONAL

> **THE HEADLINE**
>
> AI does not create new access. It makes existing exposure instant. Every finding above exists today through normal access. Point Copilot, ChatGPT Enterprise or Claude at this estate and any user reaches it in one question, at scale.

02 Readiness model, scope and context

The three stages of AI

AI adoption moves through three stages. Each adds capability and risk, and each depends on the stage below it.

STAGE	WHAT IT IS	NEW RISK
1 · Productivity AI (consume)	AI assistants over your own content: Microsoft 365 Copilot, and ChatGPT Enterprise or Claude connected to SharePoint and OneDrive.	Oversharing, unclassified sensitive data, ownerless content surfaced in AI answers.
2 · Connected AI (integrate)	AI joined to external tools and connectors, where data moves outside the tenant.	Data egress to third-party AI, shadow AI, connectors with excessive scope, loss of residency control.
3 · Agentic AI (build)	Custom agents and workflows that act on their own and call other systems.	Non-human identities with broad standing access, no MFA anchor, prompt injection, unmonitored activity.

Scope of this engagement

This assessment covers Stage 1 readiness: whether AI can be enabled safely over corporate data for productivity use across the estate. Stage 2 and Stage 3 controls depend on the specific AI tools and use cases the business intends to pursue and are scoped separately. The seven gates apply whether the AI is Microsoft 365 Copilot grounding on the tenant, or a third-party tool reaching the same content through a connector.

Access and method

Findings are based on data exported from SharePoint Online and OneDrive together with data security posture management findings. Controls that depend on Entra ID or Conditional Access could not be independently verified and are recorded on the basis of information provided by the client. Microsoft reports give a point-in-time view. Posture management adds correlations a point-in-time export cannot. At the time of writing, a first pass had scanned a fraction of the estate.

Design only

This deliverable defines the target and the sequence. Execution and remediation stay with the client, or with Cornerstone under a separate build engagement. There is no obligation to Cornerstone to act on it.

03 Engagement findings summary

Tenant overview

Three discrete Microsoft 365 tenants with no cross-tenant federation. Combined AI indexing surface is roughly 46,000 locations.

TENANT	SHAREPOINT SITES	ONEDRIVE	NOTES
Home region	~1,600	~2,600	Deepest-scanned. Highest record exposure by volume.
Region NA (largest)	~16,300	~24,000	~13,200 orphaned OneDrive (~430 TB). Highest sharing risk.
Region EU (smallest)	~30	~190	Small footprint, same absence of governance controls.

Sharing exposure

EXPOSURE	VOLUME	AUDIENCE	PRIORITY
Anyone links (unauthenticated)	~16,800	Anyone with the URL. No sign-in. Includes anonymous internet users.	CRITICAL
All-internal-users permission	~1,800 sites	Every licensed internal user, without explicit assignment.	HIGH
Orphaned OneDrive accounts	~13,200	Departed owners. Existing sharing links stay live.	HIGH
Single high-risk collaboration site	~410,000	Guest permission rows on one site. Highest single concentration.	CRITICAL
Third-party firm access	~9,500 rows	External identities scoped to specific files. Requires review and scoping.	HIGH

Privileged application risk

A third-party SharePoint management application holds tenant-wide full control across all three tenants. A single compromise of that identity bypasses every site-level control in the design. The fix is not to remove the app but to harden the identity: workload Conditional Access, credential rotation on a set cadence, restricted network egress and quarterly attestation of continued need.

Systemic governance gaps

- No consistent site ownership or data owner assignment.
- No structured group model. Direct user permission assignments are common.
- No sensitivity label coverage. Most content is unlabelled.
- No lifecycle process for inactive sites or orphaned drives.
- No external sharing governance. Guest access is uncontrolled.

04 Microsoft admin export findings

Findings from direct review of SharePoint admin centre configuration, independent of the data-level findings. Abridged here to the highest-severity items.

Critical and high configuration findings

FINDING	CURRENT STATE	AI IMPACT	RISK
Default sharing link inherits Anyone	New sites create Anyone links by default.	Every new workspace inherits exposure at creation. AI surfaces it to anyone with the URL.	CRITICAL
No sensitivity labels on groups or Teams	Container labelling absent across reviewed groups.	No container-level protection. AI cannot tell sensitive from public.	CRITICAL
Site lifecycle management off	No inactive, ownership or attestation policies.	Departed projects and dissolved teams stay indexed indefinitely.	HIGH
No domain allowlist for external sharing	Sharing permitted to any external domain.	Sensitive material can leave the tenant with no approval or detection.	HIGH
Content discovery unrestricted on sensitive sites	Restrict content discovery off on reviewed sites.	Confidential sites surface in org-wide search and AI to users with no business need.	HIGH

Compliant settings confirmed

Not everything is a gap. The following were confirmed correct at org level and require no immediate change, subject to a site-level check during remediation:

- Default link type set to specific people at org level.
- Default link permission set to view.
- Guest access expiry configured at 30 days, with re-authentication.
- Guests cannot share items they do not own.
- Custom scripts blocked at site level.

> **BALANCE**
>
> Org-level settings establish the ceiling, not the floor. Each is confirmed again at site level during the audit, because a site can override the tenant default.

05 Data security posture findings

Posture management scans data at rest across the estate and classifies it, finding correlations a configuration export cannot. Figures below are illustrative.

Finding distribution

SEVERITY	FINDINGS	RECORDS AT RISK	PRIMARY CATEGORY
CRITICAL	~60	~5,900,000	Missing labels, residency violations
HIGH	~300	Not quantified	Oversharing, unclassified PII
MEDIUM	~90	Not quantified	External access, stale permissions
LOW	~50	—	Hygiene, minor drift

Findings by category

CATEGORY	WHAT IT IS	IMPACT ON AI
Missing sensitivity labels	Largest set. Spans PII, financial records, contracts.	Purview cannot enforce protection or constrain grounding. Sensitive content is indistinguishable from public in AI answers.
Home-region residency violations	~40 findings. Home-region data held outside the region.	A potential regulatory matter requiring review before relocation.
Oversharing and access drift	Members and guests hold access beyond current role.	AI surfaces out-of-role content to those users at scale.
Stale and orphaned content	Inactive sites and departed-staff drives.	Content stays discoverable with no disposition workflow.
Insider risk indicators	Anomalous access and bulk-download signals.	Behavioural risk compounds once AI is live.

WHY THE RECORD COUNT MATTERS

The aggregate record count is the primary measure of exposure. Reduction is the explicit success metric for Phase 1 and Phase 2. Enabling AI amplifies the consequence of every record left exposed.

06 AI readiness scorecard

Seven foundational gates must be satisfied before AI can be recommended over corporate data. The gates hold whether the AI is Copilot, ChatGPT Enterprise, Claude or a connector. Each is scored on evidence. One gate is met. Six remain open, three of them critical.

G1	**Tenant and identity baseline** MFA confirmed by the client. Conditional Access review recommended.	MET
G2	**Content lifecycle and ownership** Ownerless sites and orphaned drives across the estate.	NOT MET · HIGH
G3	**Oversharing remediation** ~16,800 unauthenticated links. No domain allowlist confirmed.	NOT MET · CRITICAL
G4	**Data classification and protection** ~3 percent label coverage. No active DLP confirmed.	NOT MET · CRITICAL
G5	**Interaction guardrails and insider risk** No posture management for AI or insider risk policies scoped to AI.	NOT MET · HIGH
G6	**Regulatory, audit and residency** ~40 residency findings. Audit and eDiscovery for AI prompts not confirmed.	NOT MET · CRITICAL
G7	**Adoption, measurement and operations** No pilot cohort, scenario backlog or measurement baseline.	NOT MET · HIGH

How each gate is scored

Ratings are evidence-based. A gate cannot rate above not started without an artefact from the environment. The scale runs not started, foundational, operational, optimised. The model is re-applied on a defined cadence after rollout, quarterly for the first year, then six-monthly.

> **NOT A WALL OF RED**
>
> One gate is met and several controls were confirmed sound. The scorecard tells you where to spend effort and where you are already in good shape, not that everything is broken.

07 Design principles

Seven principles govern every design decision in the full document. They are derived from Cornerstone data management practice and adapted to the client estate.

Identity first

All access granted through security groups, never direct user assignment. Auditable and lifecycle-aligned.

Least privilege

Read is the default. Write and full control are explicit and reviewed on a cadence.

Explicit access

No site is reachable through inheritance or a stray link alone. Access is always assigned.

Tier by risk

Sites are classified into tiers that set sharing, labelling and review requirements.

Data owner accountability

Every site has a named owner accountable for content and access decisions.

Structured taxonomy

A consistent folder and naming model across governed sites. No ad hoc structure.

Safe for AI by design

The architecture is built to clear all seven gates once remediation is complete, for any AI tool.

WHY PRINCIPLES, NOT JUST FIXES

A findings list closes today's gaps. Principles stop them reopening. Every rule in the design traces back to one of these seven, so the estate stays safe for AI as it grows.

Sections 8 to 16, design detail

The full document turns these principles into a target architecture: SharePoint information architecture, folder taxonomy and metadata, the group model, the access and permission model, data ownership, external sharing, governance, and the Purview design. The next pages show the shape of that design. The controlled vocabularies, naming conventions and configuration values are provided in the engagement, against your own estate.

08 Target design: architecture, taxonomy, groups

SharePoint information architecture

A hub-and-site model organised by business unit and region. Hubs provide navigation, aggregated news and search scope. Content lives in associated member sites. Every site is classified into one of three tiers at creation.

TIER	CONTENT	SHARING RESTRICTION	REVIEW
Tier 1 · Governed	Permanent, sensitive or regulated content	No Anyone links. Authenticated users only. Guest access by approval.	Quarterly
Tier 2 · Project	Time-bound project sites, moderate sensitivity	Authenticated users. Controlled guest access by invitation.	Semi-annual
Tier 3 · Archive	Expired or read-only sites, retained for record	Read-only. No new sharing. Removed from hub navigation.	Annual

Folder taxonomy and group model

A three-layer folder model (header, sub-function, working) applies to governed sites, with access assigned by group at each layer. Four group types carry access: site permission groups, role groups, team groups and data owner groups. Site and group names follow a fixed pattern by region, business unit and access level.

ABRIDGED IN THIS SAMPLE

The controlled vocabulary, the full naming conventions for sites and groups, the metadata column set, and the attribute-based membership rules are provided in the engagement. They are the reusable part of the intellectual property and are tailored to your estate.

ELEMENT	PATTERN (SHAPE ONLY)
Site name	[Region]-[Business unit]-[Site]
Site permission group	[Prefix]-[Region]-[Business unit]-[Site]-[Access]
Data owner group	[Prefix]-[Region]-[Site]

09 Access, ownership, sharing and governance

Access model

Three permission levels, granted only through groups. Groups nest so that access is managed in one place and follows the joiner, mover, leaver lifecycle.

LEVEL	GRANTED TO	BASIS
Read	Default for members via role or team group	Standing, reviewed on cadence
Contribute	Working members of a site	By data owner approval
Full control	Site owners only, by exception	Named, attested quarterly

Data ownership

Every site has a named data owner, accountable for content, access decisions and quarterly review. Ownership is tracked centrally through a dedicated group, not held in a person's head. Access reviews run quarterly for Tier 1, semi-annually for Tier 2.

External sharing, target state

Anyone links are removed and replaced with specific-person links. A domain allowlist limits external sharing to approved partners. Guest accounts carry expiry and periodic re-attestation. External sharing capability is restricted to an approved security group.

Governance model

A standing operating model keeps the estate safe after remediation: defined governance roles, a site lifecycle from request to archive, a disposition workflow for orphaned drives, and a governance calendar of recurring reviews. Policy is monitored continuously by posture management, with drift routed back to the owner.

> **THE POINT**
>
> Access control, ownership, sharing and governance are one system. Fix them together and AI inherits a clean estate. Fix them in isolation and exposure returns within a quarter.

10 Microsoft native capability and posture management

Microsoft native tooling enforces. Posture management discovers and monitors at a breadth native tooling does not reach. The design uses both, each for what it does best.

OUTCOME	MICROSOFT NATIVE	POSTURE MANAGEMENT (DSPM)
Discovery across the estate	Within Microsoft 365	Across clouds, databases and file shares too
Classification at scale	Label-driven, needs coverage	Automated classification of unlabelled data
Enforcement for AI	Labels, DLP, Copilot controls	Feeds classification into native enforcement
Drift monitoring	Point-in-time	Continuous, with alerting

Purview readiness design

The full document specifies the label taxonomy, auto-labelling logic, retention and a data loss prevention design that covers AI locations, sequenced so labels reach coverage before enforcement is switched on. Labels are what let any AI honour the client's protection model, so this design is the hinge of the whole programme.

ABRIDGED IN THIS SAMPLE

The label taxonomy, auto-labelling conditions, retention schedule and the data loss prevention rule set are provided in the engagement. Published verbatim they would be a copy-and-run configuration, so they are withheld here.

11 Readiness pathway and AI risk scenarios

Gate-clearing pathway

The gates clear in a set order. Identity and residency first, then oversharing and lifecycle, then classification and data loss prevention, then guardrails and operations. Enabling AI, of any kind, follows the clearance of the critical and high gates.

AI risk scenarios

STAGE	SCENARIO	PREVENTED BY
1 · Productivity	A user asks an AI tool to summarise a topic. A broadly shared confidential document appears. The exposure existed already; AI made it instant.	Access control, labels, DLP (G3, G4)
1 · Productivity	An AI answer surfaces personal or financial data from an ownerless site with no label.	Classification, DLP, ownership (G2, G4)
2 · Connected	Staff paste client data into a public AI tool, or a connector syncs content to a third-party AI outside residency.	Shadow AI discovery, guardrails, DLP to AI (G5)
3 · Agentic	An over-permissioned agent reaches and moves data at machine speed with no user to challenge it.	Non-human identity governance (forward scope)

INFORMED RISK ACCEPTANCE

Where the business chooses to proceed ahead of a gate, the residual risk is named, owned and time-bound. The report does not hide a decision to accept risk. It documents it.

12 Configuration reference and remediation procedure

The full document carries the settings reference and the step-by-step remediation the client's administrators execute. Sections 19 and 20 are the runnable core and are the most heavily abridged in this sample.

What these sections contain

- Organisation, site and document-library configuration reference, with the target value for each setting by tier.
- AI and Copilot exclusion settings for sites that must never be indexed.
- Bulk-change scripting to apply settings across thousands of sites.
- A seven-step external sharing remediation procedure, phased by risk.

A single illustrative line, to show the format:

SETTING	TARGET (TIER 1)	WHERE
Default sharing link type	Specific people	Site sharing settings

WITHHELD FROM THIS SAMPLE

The full configuration reference, the AI and Copilot exclusion settings, and the bulk-change scripts are provided in the engagement. This is the material a client can run directly, so it is not published in a sample. Without it the findings and roadmap still make the value clear; with it, the sample would be a free runbook.

13 Remediation roadmap and responsibilities

The fix is sequenced into four phases across a year. Each phase maps to the gates it clears.

DAYS 0-30

Critical risk remediation

Revoke or convert unauthenticated links on the highest-risk sites. Lock down org-level sharing. Harden the privileged management identity. Begin residency remediation. Clears the path on G3 and G6.

DAYS 30-90

Governance foundation

Assign owners and the data-owner model. Deploy the label taxonomy and auto-labelling. Introduce guest expiry and site lifecycle policies. Builds G2 and G4.

DAYS 90-180

Architecture and compliance

Roll out the target SharePoint and access architecture. Activate data loss prevention once label coverage is sufficient. Stand up insider risk and posture management for AI. Builds G4 and G5.

DAYS 180-365

Adoption and operations

Pilot AI with a defined cohort and scenario backlog. Establish measurement, an AI council and a reassessment cadence. Builds G7.

Responsibilities (RACI, summarised)

ACTIVITY	CORNERSTONE	CLIENT IT	CLIENT BUSINESS	TOOLING
Assessment and design	R / A	C	I	C
Remediation execution	C	R / A	C	R
Data owner decisions	I	C	R / A	I
Ongoing monitoring	I	A	I	R

14 Risks, training and appendices

Selected risks and dependencies

ITEM	NOTE	LIKELIHOOD
Link revocation at scale	Bulk revocation can disrupt business-critical sharing. Analyse and communicate first.	HIGH
Residency remediation	May carry legal and contractual implications. Engage legal before moving data.	MEDIUM
Privileged app dependency	Management app may run scheduled jobs. Confirm minimum permissions before change.	HIGH
Attribute data quality	Inconsistent directory attributes reduce automated group accuracy. Audit before build.	MEDIUM

Training plan

Adoption depends on people, not only controls. The full plan covers governance fundamentals for all staff, data owner responsibilities, administrator workshops, posture management operations, and safe AI use for the pilot cohort, each tied to the phase it supports.

Appendices

A: regional data summary. B: AI readiness checklist. C and D: naming references. E: the seven-gate model. F: glossary. G: the stage model. The seven-gate model in Appendix E is the assessment method itself, reproduced below.

GATE	READINESS DOMAIN	PURPOSE
G1	Tenant and identity baseline	Licensing, identity and network conditions AI needs to operate in your risk appetite.
G2	Content lifecycle and ownership	Current, owned and governed content before any AI indexes it.
G3	Oversharing remediation	Reduce the surface area AI can ground on.
G4	Data classification and protection	Labels, encryption and DLP so AI honours your protection model.
G5	Interaction guardrails and insider risk	AI interactions inherit protection and are observed for risky use.
G6	Regulatory, audit and residency	AI operates inside regulatory, evidentiary and residency commitments.
G7	Adoption, measurement and operations	The operating model to turn a rollout into sustained value.

15 What the full deliverable adds

This sample reproduces the findings and the shape of the design and roadmap. The detail that lets your team build is delivered in your own engagement, against your own data. Held back here:

Held back from this sample

- Target SharePoint information architecture in full: hub-and-site model, tiering and naming vocabulary.
- Group architecture and naming conventions for access at scale.
- The three-layer folder taxonomy and metadata model.
- Purview label taxonomy, auto-labelling logic and data loss prevention design.
- SharePoint configuration reference, AI exclusion settings and bulk-change scripting.
- Step-by-step external sharing remediation procedure.
- Full RACI matrix, risk register, assumptions, dependencies and training plan.
- Per-tenant regional data summaries with the real figures.

Why the difference matters

A findings list tells you that you have a problem. A design tells you exactly how to fix it, in your environment, in the right order, with the controls named and the owners assigned. It applies whether the AI you enable is Microsoft 365 Copilot, ChatGPT Enterprise, Claude, Gemini or a connector that syncs your data. The sample proves the output is clear and actionable. The engagement gives you the build.

> **NEXT STEP**
>
> Book an AI readiness assessment and we will score your seven gates against your own tenant, then hand you a report like this one, unredacted, for your estate. Visit cornerstonecyber.com.au or contact the team.

This document is an anonymised, illustrative sample. It does not describe any identifiable organisation. Figures have been altered and are not real measurements. Prepared by Cornerstone Cyber.