TinyResearch Sample Report: Enterprise Data Protection

TinyResearch · Sample research report · DRACO

This sample presents a complete TinyResearch report in response to the question below. The report text is reproduced unchanged, with its cited sources listed at the end.

Original research question

Data Loss Prevention for ChatGPT Enterprise: A Comprehensive Decision Memo

- Executive Summary
- Part I: Foundational Architecture — What Is and Is Not Supported
  - 1.1 The Capability Gap in Black and White
  - 1.2 Why the Architecture Works This Way
- Part II: Option-by-Option Evaluation
  - Option 1: Purview Compliance API Connector for ChatGPT Enterprise (Post-Hoc Governance)
    - What It Does
    - What It Does Not Do
    - Setup Requirements
    - Licensing
    - Known Limitations
  - Option 2: Microsoft Copilot for Microsoft 365 as an Alternative to ChatGPT Enterprise
    - Native Purview DLP Coverage
    - Comparison with ChatGPT Enterprise
    - Critical Caveat for Copilot
    - Verdict on Option 2
  - Option 3: Microsoft Defender for Cloud Apps (MCAS) / CASB and Proxy-Based Controls

What MCAS Can Do for ChatGPT Enterprise

Exact Implementation Steps for MCAS Session Control on ChatGPT Enterprise

MCAS Known Limitations for ChatGPT Enterprise

MCAS vs. Purview Endpoint DLP Priority

Option 4: ChatGPT Enterprise Native Admin Controls and API-Side Controls

What OpenAI Provides Natively

Part III: Third-Party DLP Tools for ChatGPT Enterprise

Microsoft-Listed Compliance API Partners

Key Third-Party Tools in Detail

Part IV: Layered DLP Architecture — Recommended Corporate Implementation

Strategic Framework

Concrete Implementation Steps

Part V: Licensing Summary

Part VI: Trade-Offs and Decision Framework

Scenario-Based Recommendations

Known Limitations Summary (Consolidated)

Part VII: Forward-Looking Considerations

Conclusion and Final Recommendations

Source links

Original research question

Compare the available options for achieving true Data Loss Prevention (DLP) for ChatGPT Enterprise when integrating with Microsoft Purview. Specifically evaluate: 1) whether ChatGPT Enterprise has native Purview DLP support, 2) using Copilot for Microsoft 365 (with Purview) as an alternative, 3) Microsoft Defender for Cloud Apps / CASB and proxy-based controls, and 4) ChatGPT Enterprise admin/settings and API-side controls. Recommend the best option(s) for a corporate environment, list concrete implementation steps, trade-offs, and known limitations.

Data Loss Prevention for ChatGPT Enterprise: A Comprehensive Decision Memo

Executive Summary

The central finding for decision-makers: Microsoft Purview does **not** provide native, real-time DLP enforcement inside ChatGPT Enterprise. What Purview *does* provide through its Compliance API connector is a robust **post-hoc governance layer**—audit, eDiscovery, retention, Communication Compliance, Insider Risk Management, and data classification—but with a critical, documented gap: the DLP, Sensitivity Labels, and Encryption rows in Microsoft's own capability table carry explicit ✕ marks for ChatGPT Enterprise [1][7][35]. This is the single most important fact a CISO or compliance officer must internalize before designing a control architecture.

True DLP enforcement—blocking or redacting sensitive data *before* it reaches OpenAI's infrastructure—must be achieved through **edge-plane controls**: Endpoint DLP on managed devices, Browser DLP in Edge for Business, or network/CASB-layer inspection via Microsoft Defender for Cloud Apps or a third-party SSE/SASE partner. A fully layered posture for a corporate environment therefore requires combining Purview's post-hoc governance with one or more of these edge enforcement planes, and optionally augmenting with third-party AI-native DLP tools.

Microsoft Copilot for Microsoft 365, by contrast, receives full inline Purview DLP enforcement natively and represents a fundamentally different (and substantially more controllable) compliance posture than ChatGPT Enterprise [5][7][12][16].

Part I: Foundational Architecture — What Is and Is Not Supported

1.1 The Capability Gap in Black and White

Microsoft's official documentation for managing ChatGPT Enterprise with Purview includes an explicit capability matrix [1][7][35]:

Microsoft Purview Capability	ChatGPT Enterprise	Microsoft 365 Copilot
Data Loss Prevention (DLP)	✕ Not supported	✓ Full inline enforcement
Sensitivity Labels	✕ Not supported	✓ Native enforcement
Encryption without labels	✕ Not supported	✓ Supported
Data Classification (SITs/classifiers)	✓ Supported	✓ Supported
Auditing / Unified Audit Log	✓ Supported	✓ Supported
Insider Risk Management	✓ Supported	✓ Supported
Communication Compliance	✓ Supported	✓ Supported
eDiscovery	✓ Supported	✓ Supported
Data Lifecycle Management (Retention)	✓ Supported	✓ Supported
DSPM for AI reporting	✓ Supported	✓ Supported
Compliance Manager	✓ Supported	✓ Supported

The implication is architectural: the Purview ChatGPT Enterprise connector is a **camera, not a lock** [7]. It records everything after it happens and prevents nothing at the platform level.

1.2 Why the Architecture Works This Way

The Purview connector calls OpenAI's Compliance API on a scheduled basis, pulling interactions that have already occurred and landing them inside Purview's compliant storage [7]. Nothing in that data path sits between the user and the model. Critically, because of OpenAI's current API limitations, **conversations are ingested with approximately a 24-hour delay** [2][7]. This is journaling for the AI era—retrospective, not preventive [7].

By contrast, Microsoft 365 Copilot is built inside Microsoft's own service boundary. Purview's DLP engine sits inline, evaluating prompts before Copilot processes them and blocking responses when policy conditions are met [5][7] [27]. This architectural difference—internal versus external AI service—explains the entire gap.

Part II: Option-by-Option Evaluation

Option 1: Purview Compliance API Connector for ChatGPT Enterprise (Post-Hoc Governance)

What It Does

The ChatGPT Enterprise connector for Purview, generally available since December 2025 [40], enables organizations to:

- **Archive and retain** ChatGPT Enterprise prompts and responses. Interactions are stored in users' Exchange mailboxes, making them subject to standard retention policies by selecting the "Enterprise AI apps" location [1][44].
- **Audit** all user interactions in the Unified Audit Log, capturing timing, user identity, Microsoft 365 services accessed, and sensitivity labels on referenced files [1][44].
- **eDiscovery**: Because interactions land in Exchange mailboxes, legal hold, search, preservation, and export are fully functional. Searches use the `ItemClass` property with `IPM.SkypeTeams.Message.ConnectedAIApp.Connector.<ChatGPTEnterprise>` for local machine activities or `IPM.SkypeTeams.Message.CloudAIApp.SaaS.<AppID>` for browser-based activities [1].
- **Communication Compliance**: Detect regulatory violations including hate speech, corporate sabotage, credential sharing, jailbreak attempts, and prompt injection attacks using ML models and Sensitive Information Types [3][4][1].
- **Insider Risk Management**: The "Risky AI usage" policy template detects prompt injection attacks and access to protected materials; Adaptive Protection can dynamically escalate DLP enforcement for elevated-risk users [1][4][48].
- **Data classification**: Sensitive Information Types (SITs) and trainable classifiers run against ingested prompts and responses to surface sensitive content in the DSPM for AI dashboard [1][4].
- **DSPM for AI reporting**: Metrics include total interactions over time, sensitive interactions per AI app, Insider Risk severity per user and app [1][39].

What It Does Not Do

It does **not** block, warn, or redact any prompt or response in real time. It does not apply sensitivity labels to ChatGPT outputs. It does not enforce encryption [1][7][35]. A user can paste an entire customer database into ChatGPT Enterprise and the connector will faithfully record that it happened—after the fact [7][35].

Setup Requirements

1. Enable **pay-as-you-go billing** linked to an Azure subscription (required as of May 1, 2025; metered per prompt/response) [4][7][39].
2. Request a **ChatGPT Enterprise Compliance API key** from OpenAI support via email (no self-service toggle exists; specify the key name, last four digits, and desired scope—read, delete, or both) [7].
3. Store the API key in **Azure Key Vault** [4][47].
4. Assign **Graph API roles** to the managed identity: `Purview.ProcessConversationMessages.All` (GUID: `a4543e1f-6e5d-4ec9-a54a-f3b8c156163f`) and `User.Read.All` (GUID: `df021288-bdef-4463-88db-98f22de89214`) via Cloud Shell [4][47].
5. Add ChatGPT Enterprise as a **data source in Purview Data Map** and configure connector scans [47].
6. Create a **collection policy** (via DSPM for AI one-click or manually via Purview DLP) to authorize ingestion of prompts and responses—required since May 19, 2025 [4][2].
7. Run initial full scan (may exceed 24 hours depending on data volume) [4][2].

Licensing

- Microsoft 365 E5, or the Microsoft Purview Suite add-on ($12/user/month) added to E3 [39][52].
- Pay-as-you-go metered charges on Azure for each prompt/response processed [4][7][39].
- A Purview Enterprise account (tied to Azure subscription) is required, not a standard M365-only tenant [47].

Known Limitations

- **24-hour ingestion delay** makes real-time response impossible [2][7].
- No DLP, label, or encryption enforcement [1][7][35].
- Connector does not support scoped scans, labeling, access policies, lineage, data sharing, or live view [2].
- Permissions configuration is error-prone and silent—misconfiguration produces no error messages; data simply does not appear in DSPM for AI [47].
- Pay-as-you-go billing is consumption-based and not included in E5 flat licensing [7][39].

Option 2: Microsoft Copilot for Microsoft 365 as an Alternative to ChatGPT Enterprise

Native Purview DLP Coverage

Microsoft 365 Copilot operates inside the Microsoft 365 service boundary and integrates with Purview DLP through a dedicated policy location—"Microsoft 365 Copilot and Copilot Chat"—available only in the Custom policy template [5]. This enables four categories of inline, real-time enforcement:

1. **Block external web search grounding** when a prompt contains configured SITs (credit card numbers, passport numbers, SSNs, or custom SITs). Copilot continues generating responses from permitted internal M365 data sources [5].
2. **Block Copilot from processing sensitive prompts entirely**, preventing any response when prompts contain sensitive data (in preview, rolling out broadly) [5][35].
3. **Restrict Copilot from using files and emails with specific sensitivity labels** as grounding data [5][16][35].
4. **Block external email grounding** (in preview): prevent Copilot from using emails received from external domains as grounding data [5].

Additionally, Copilot interactions are audited natively, subject to retention policies, searchable via eDiscovery, monitored by Communication Compliance (with Prompt Shields for adversarial input detection and Protected Materials for copyright detection [21]), and governed by Insider Risk Management with the Risky AI usage template [1][44][21].

Comparison with ChatGPT Enterprise

Dimension	Microsoft 365 Copilot	ChatGPT Enterprise
Inline DLP enforcement	✓ Full (blocks prompts, grounding, web search)	✕ None natively
Sensitivity label enforcement	✓ Native inheritance and propagation	✕ Not supported
Data residency	✓ Stays in M365 tenant; EU Data Boundary commitment [12]	Partial (10 global regions; EU commitment since 2024 [42])
eDiscovery	✓ Full	✓ Full (via connector)
Communication Compliance	✓ Full with Prompt Shields	✓ Full (via connector, 24h delay)
Insider Risk Management	✓ Native	✓ Supported (via connector)
Data leaves Microsoft infrastructure	No	Yes (passes through OpenAI's infrastructure) [11]
Model training on data	No (by policy) [12]	No (by contract) [13][19]
Oversharing risk	High (surfaces existing M365 permissions) [11][23]	Lower for internal data; higher for external paste/upload
Licensing	M365 E3/E5 + Copilot license ($30/user/month) [39]	Negotiated enterprise contract (~150 seat minimum) [42]

Critical Caveat for Copilot

Copilot's primary security risk is **permissions oversharing**, not data exfiltration to an external provider. Copilot can access any file the user has permission to access in M365, and if SharePoint/OneDrive ACLs are misconfigured—which is common in most tenants—Copilot can synthesize and surface sensitive content the user technically can see but should not see in context [11][23]. Microsoft's

SharePoint Advanced Management and Purview's data risk assessments are prerequisites before activating Copilot at scale [23][39].

Additionally, Purview DLP for Copilot has documented technical limitations:

- DLP cannot scan the contents of files **uploaded directly into prompts**; only typed text is evaluated [5][35].
- SITs created through **document fingerprinting** are not supported [23].
- Policy changes take **up to four hours** to propagate [5][35].
- The `Content contains SITs` and `Content contains sensitivity labels` conditions **cannot be combined** in the same rule [23].

Verdict on Option 2

For organizations seeking the strongest available Purview DLP coverage on an AI assistant, Microsoft 365 Copilot is unambiguously superior to ChatGPT Enterprise from a governance standpoint. However, it does not eliminate the ChatGPT Enterprise problem if the organization is running both (as most Fortune 1000 enterprises are projected to do by end of 2026 [11]). The two tools have different leak profiles: Copilot leaks are primarily internal (permissions exposure), while ChatGPT Enterprise leaks are primarily external (prompt paste and file upload) [11].

Option 3: Microsoft Defender for Cloud Apps (MCAS) / CASB and Proxy-Based Controls

What MCAS Can Do for ChatGPT Enterprise

Microsoft Defender for Cloud Apps provides two enforcement mechanisms relevant to ChatGPT Enterprise:

A. Shadow IT Discovery and App Sanctioning

MCAS maintains a catalog of 31,000+ SaaS applications [29] and includes a built-in Generative AI category automatically grouping 1,200+ known AI services including ChatGPT, Gemini, and DeepSeek [14]. Organizations can:

- Mark ChatGPT.com as **unsanctioned** and block access via Microsoft Defender for Endpoint's Network Protection (blocking occurs at the network level, regardless of browser or HTTP client) [14].
- Simultaneously mark the ChatGPT Enterprise workspace URL as **sanctioned** to permit corporate access while blocking personal accounts.
- Generate Shadow IT reports to discover employees using unauthorized AI tools.

B. Conditional Access App Control (Session Policies)

MCAS can deploy reverse-proxy session policies for web applications. For ChatGPT Enterprise specifically:

- Session policies can **monitor all activity** in the ChatGPT web session.
- Policies can **block file uploads** (with or without content inspection) based on file metadata, sensitivity label, file name, extension, or size.
- Policies can **block file downloads** from ChatGPT Enterprise.
- Files being downloaded through the session can have **Purview sensitivity labels applied automatically** (for Office and PDF file types only) [8][26].
- Content inspection can scan files against sensitive data patterns (e.g., social security numbers) before allowing or blocking upload [8][28].

Exact Implementation Steps for MCAS Session Control on ChatGPT Enterprise

1. **Verify licensing**: Defender for Cloud Apps requires Microsoft 365 E5, EMS E5, the Defender or Purview suites, E5 Information Protection and Governance, or a standalone MDCA license. Conditional Access App Control additionally requires **Entra ID P1 per user** [26][41].
2. **Configure Conditional Access policy in Entra ID**: Create a CA policy targeting ChatGPT Enterprise users, with session control set to "Use Conditional Access App Control" → "Use custom policy." The CA policy must explicitly include the App Control session control; omitting it prevents any routing to MCAS [41].
3. **Onboard ChatGPT.com to Conditional Access App Control** in the Defender for Cloud Apps portal under Connected Apps.

4. **Create a Session Policy** in the Microsoft Defender portal:
    - Policy type: Session policy
    - Session control type: "Control file upload (with inspection)" or "Monitor only"
    - Configure file filters: Sensitivity label, file name, extension, file size
    - Enable content inspection using Data Classification Service
    - Set action: Block, with optional user notification
5. **Test with a pilot group** before broad deployment.
6. **Layer with Endpoint DLP** to handle the Teams desktop client and native apps (which bypass the MCAS reverse proxy) [26][41].

MCAS Known Limitations for ChatGPT Enterprise

These limitations are officially documented and material:

- **Session controls apply only to web browser–based access**; the Teams desktop client, native ChatGPT desktop app, and non-browser HTTP clients are not covered [26][38].
- **File size ceiling**: Session policies apply to files up to **50 MB**; larger files follow tenant-level allow/block settings regardless of policy [38][26].
- **Content inspection ceiling**: Only files smaller than **30 MB** with fewer than **1 million characters** are scanned; larger files follow the "Always apply" fallback action [38][26].
- **Files encrypted with sensitivity labels cannot be read during content inspection**; policy falls back to the "always apply" action [38].
- **IPv6 requests bypass IP-based policy rules** entirely [38][26].
- **Edge in-browser protection** (an alternative to reverse proxy for MCAS): requires the last two stable versions of Edge for Business, commercial tenants only, and does not support Google Workspace [26].
- **Sensitivity label enforcement** via Protect action works only on Office formats and PDF; the original file in the cloud app is not modified—only the

downloaded copy is labeled [26].

- **Non-interactive token flows** (e.g., Teams→SharePoint) bypass session controls because users are not re-authenticated [26].
- **No "block with override" option**: Session controls offer only audit, block, or protect—no justified-override workflow as DLP provides [26].
- **Cached policy data** in Edge can persist session enforcement even after users are removed from the CA policy; the documented fix requires manual deletion of `mda_store.1.txt` from the user's Edge profile [26].
- **File policies in Defender for Cloud Apps** will be retired on **January 6, 2027**; organizations should migrate to Purview DLP or auto-labeling policies [28].
- **Google Workspace is not supported** with Edge in-browser protection for real-time DLP scans [26].
- **B2B guest users in Teams** are not protected by session policies [38][26].
- **Drag-and-drop uploads** of mixed content (files and folders) can trigger over-blocking of the entire selection [38].

MCAS vs. Purview Endpoint DLP Priority

When the same context and action are configured in both Endpoint DLP and a Defender for Cloud Apps session policy, **Endpoint DLP takes precedence** [26]. This determines design priority in layered architectures.

Option 4: ChatGPT Enterprise Native Admin Controls and API-Side Controls

What OpenAI Provides Natively

ChatGPT Enterprise offers a substantial set of administrative controls that, while they do not constitute DLP in the technical sense, significantly reduce the attack surface:

Identity and Access Management

- **SAML SSO** with Okta, Azure AD, Google Workspace, Ping, and OneLogin [42] [17].
- **SCIM provisioning** (via WorkOS) supporting automated user lifecycle management with Okta, Entra ID, Google Workspace, and OneLogin—assigning/removing a directory user automatically invites/deactivates the ChatGPT account [10][17].
- **IP allowlisting** to restrict workspace access to corporate network/VPN IP ranges, with a separate Compliance API IP allowlist [17][18].
- **RBAC** with custom roles and granular feature permissions (agent mode, Canvas, Codex, GPT creation, memory, tool use) assignable to user groups [17].

Data Controls

- **No training on customer data** by default—contractual commitment, not an opt-out [19][13][42].
- **Customer-managed encryption keys (EKM)** announced December 2025: organizations can use their own cloud KMS (AWS, Azure, GCP) to encrypt all customer content; even OpenAI cannot decrypt without the enterprise's permission [17][42].
- **Data residency**: 10 global regions available; EU data at-rest commitment since 2024 for EU-contracted entities [17][42].
- **Custom data retention windows**: Admins configure chat history and memory retention (number of days or never); default is 30 days [42][17].
- **Conversation controls**: Memory is disabled by default; workspace admins can configure it at workspace level independently of individual user settings (important audit point: verify workspace-level toggle separately) [18].

Compliance API and Audit

- **OpenAI Compliance Logs Platform**: Immutable, append-only JSONL log files with minutes-level latency, covering conversation message logs, Admin Audit logs, User Authentication logs, and Codex Usage logs [10][18].
- **30-day retention on the Compliance Logs Platform**; organizations requiring longer retention must implement their own continuous download and retention

pipeline [6].

- **Two access patterns**: Compliance Logs Platform (immutable, event-based) and Stateful Compliance API (query state at time of request) [6].
- **SIEM integration**: Ready-to-use download scripts for Unix/bash and Windows PowerShell provided by OpenAI for Splunk, Datadog, Microsoft Sentinel, etc. [18][42].
- Compliance API key requires email request to `support@openai.com` with no self-service scope toggle [7].

App and Connector Controls

- Apps (connectors to Google Drive, Outlook, Teams, SharePoint, Slack, etc.) are **disabled by default** in Enterprise (unlike Business tier where they are enabled by default) [18].
- Third-party GPTs are blocked by default; publishing externally is disabled by default [18].
- Custom GPT creation enabled for personal use by default; workspace and external publishing blocked by default [18].
- Admins can request **domain blocklists** for Web Browsing Agent mode [17].
- **Agent mode** (highest prompt injection surface) is enabled by default; admins should review and restrict via RBAC for sensitive teams [18].
- **Codex** supports admin-enforced `requirements.toml` with approval policies, sandbox modes, and MCP server allowlists that users cannot override [18].

What API-Side Controls Do Not Provide

OpenAI's native controls do **not** inspect prompt content for sensitive data, do not apply sensitivity labels, and do not enforce DLP policies on outbound data [11]. The commitment that "data is not used for training" is a contractual assurance about model training, not a technical control preventing data from transiting OpenAI's infrastructure [11]. Data still passes through OpenAI's systems, is logged in the customer's tenant, and can be exposed through prompt paste, connector configurations, or custom GPT actions [11].

Part III: Third-Party DLP Tools for ChatGPT Enterprise

When Purview's capabilities are insufficient and MCAS session controls carry too many limitations, the following purpose-built tools fill specific gaps. All of these integrate with ChatGPT Enterprise via the OpenAI Compliance API, endpoint agents, or browser extensions.

Microsoft-Listed Compliance API Partners

OpenAI's official Compliance Platform lists 13 compliance API integrations [18][6]:

Vendor	Integration Type	Capabilities
Microsoft Purview	Compliance API + connector	Post-hoc governance (audit, eDiscovery, IRM, Communication Compliance, classification)
Cyberhaven	Compliance API + endpoint agent	Data lineage, classification, incident response; monitors conversations, file uploads, tool invocations, custom GPT configs [36][37]
Netskope	Compliance API + SSE proxy	AI Gateway: inline prompt inspection, tenant differentiation (personal vs. corporate accounts), response inspection [46][50]
Theta Lake	Compliance API	Regulatory archiving and supervision
Zenity	Compliance API	Agentic AI security
TrojAI	Compliance API	Prompt injection and safety monitoring
Relativity	Compliance API	eDiscovery-focused
Tenable	Compliance API	Security posture
Varonis	Compliance API	Data access governance
Palo Alto Networks (Prisma CASB)	Inline CASB + NGFW	Inline DLP with App-ID for ChatGPT traffic [34]
Zscaler (ZIA)	SSE proxy	Inline DLP through Zero Trust Exchange; requires WebSocket inspection enabled by Support [13][52]
Strac	Compliance API (Claude) +	Browser DLP, endpoint DLP, MCP DLP [53]

Vendor	Integration Type	Capabilities
	browser/endpoint	

Key Third-Party Tools in Detail

Cyberhaven [36][37][51]

- Uses ChatGPT Enterprise Compliance API to monitor conversations, file uploads, tool invocations, custom GPT knowledge files, and metadata (data at rest in the AI tool, not just in motion).
- Integrates with endpoint agent and browser extension for data in motion.
- Claims 90% reduction in false positives vs. pattern-matching DLP [51].
- Full data lineage tracking: follows data from origin through AI tool destinations across endpoints, cloud, and web from a single architecture [51].
- GA since February 2026; Claude Enterprise integration expected August 2026 [36][51].
- Setup: configure in Cyberhaven Console using OpenAI API key with Compliance API read scope and ChatGPT Enterprise workspace ID [36].

Netskope [46][50][25]

- AI Gateway controls what data enters ChatGPT, Copilot, Gemini, and other AI tools and inspects what they return.
- Distinguishes personal from corporate accounts of the same AI tool to prevent personal-account bypass [25][50].
- Cloud-proxy model; AI traffic is inspected after routing through Netskope's cloud (creates latency and data-residency considerations from backhauling) [50].
- Integrated with Microsoft Global Secure Access (GSA) for coordinated TLS inspection, ATP, and DLP policies [9].
- Listed as Compliance API partner; also integrates at network/SSE layer [6] [46].

Nightfall AI [24][51][50]

- Chrome browser plugin + endpoint agents; no proxy required.
- Detects 100+ sensitive data types (PII, PCI, PHI, secrets, credentials) in real time before prompt submission [24].
- ML-based classifiers; claims 2× precision and 4× fewer false positives than legacy DLP [51].
- Supports automated redaction, user coaching, and workflow integrations [24].
- Browser-plugin model is strongest in browser; less mature for native desktop apps and non-browser paths [50][51].
- Endpoint/browser coverage added in 2025 (originally built API-first for SaaS) [51].

Zscaler (ZIA Data Protection) [13][50][52]

- Inline DLP and AI app controls through Zscaler Internet Access Zero Trust Exchange.
- Tenant profiles for ChatGPT and Claude for tenant differentiation [52].
- Requires WebSocket inspection to be enabled by Zscaler Support for full ChatGPT coverage [52].
- Cloud backhaul adds network latency; enterprise pricing and setup complexity [13].
- Deepest controls apply when AI interactions flow through browser or web-proxied channels; gaps for native desktop apps [50].

dope.security [13][50][25]

- On-device prompt inspection using LLM classification (no backhaul to data center) [13][50].
- Cloud Application Control to allow enterprise ChatGPT tenant while blocking personal logins [13].
- Claims up to 4× faster than legacy proxies due to no backhauling [13].
- Uses OpenAI zero-data-retention APIs for classification (US Patent no. 12,464,023) [13].

Palo Alto Networks Enterprise DLP [34]

- Inline content inspection via Prisma Access CASB and NGFW.
- Identifies ChatGPT traffic via App-ID `openai-chatgpt`.
- Can block or allow access to ChatGPT while selectively preventing sensitive data exfiltration through security policy rules [34].

Harmonic Security [18]

- Referenced in OpenAI guidance; AI-native DLP posture for ChatGPT Enterprise [18].

Part IV: Layered DLP Architecture — Recommended Corporate Implementation

Strategic Framework

No single control achieves complete DLP for ChatGPT Enterprise. The recommended architecture uses **three enforcement planes** plus **one governance plane**:

```
Enforcement Planes (preventive, real-time):
  1. Endpoint DLP (paste, upload, file copy on managed Windows/macOS)
  2. Browser DLP (typed text, paste in Edge; extension for Chrome/Firefox)
  3. Network/CASB DLP (proxy-based; covers web traffic broadly)

Governance Plane (detective, post-hoc):
  4. Purview Compliance API Connector (audit, eDiscovery, IRM, Communication Compliance)
```

Concrete Implementation Steps

Phase 1: Foundation and Discovery (Weeks 1–2)

1. **Inventory AI usage**: Use Defender for Cloud Apps Shadow IT discovery to identify all AI tools in use across the organization. The Generative AI category

automatically groups 1,200+ AI services [14]. Generate a report of unsanctioned AI tool usage.

2. **Classify data**: Audit sensitivity label coverage across SharePoint, OneDrive, and Exchange. Run Purview's data risk assessment for the top 100 SharePoint sites [39][43]. Classify sensitive data using SITs and trainable classifiers before deploying AI controls.
3. **Establish ChatGPT Enterprise as the sanctioned path**: Mark the ChatGPT Enterprise workspace URL as sanctioned in MCAS; mark ChatGPT.com consumer as unsanctioned. This creates a single approved access path [13][14].
4. **Configure SSO/SCIM**: Integrate ChatGPT Enterprise with Microsoft Entra ID (Azure AD) for SAML SSO and SCIM provisioning. Enable IP allowlisting for corporate network/VPN access [17][18].
5. **Audit OpenAI native settings**: Verify memory is disabled at workspace level, apps/connectors are disabled by default, agent mode is restricted via RBAC to approved groups, and Codex `requirements.toml` is configured [18].

Phase 2: Endpoint and Browser DLP (Weeks 2–4)

6. **Onboard Windows devices to Purview Endpoint DLP** via Intune/MDE. If Defender for Endpoint is already deployed, the onboarding package is shared automatically [49]. Target: all managed Windows 10/11 and macOS devices.
7. **Configure Endpoint DLP policy for AI sites** (in audit mode first):
   - Create policy using Custom template.
   - Locations: Devices.
   - Conditions: Content contains SITs (credit card, SSN, passport, PHI, custom IP classifiers) OR Content has sensitivity label (Confidential, Highly Confidential).
   - Service domains: Add `chatgpt.com`, `chat.openai.com` to the Generative AI websites service domain category.
   - Actions: Audit → progress to Block with override → Block.

- Activities: Upload to restricted cloud service domain AND Paste to browser [15][27][45][49].

8. **Configure Browser DLP for Edge for Business**:
    - Require corporate account sign-in via Intune Settings Catalog policy [49].
    - Configure Inline web traffic DLP policy with Edge for Business enforcement [49][27].
    - Enable unmanaged browser blocking (Chrome, Firefox) via Unallowed Browsers list for ChatGPT domains; deploy Purview browser extension via Intune for Chrome and Firefox [15][49].
9. **Create custom SITs for source code** (keyword-based: function, def, class, import, SELECT, INSERT, etc.) since no built-in type exists [49].
10. **Monitor for 7–10 days in audit mode** before switching to block. Review Activity Explorer for false positives; tune SIT confidence levels [27][39].

Phase 3: MCAS Session Controls (Weeks 3–5)

11. **Create Conditional Access policy** in Entra ID: scope to ChatGPT Enterprise users, session control = "Use Conditional Access App Control" → custom policy [41].
12. **Onboard ChatGPT.com to Conditional Access App Control** in Defender for Cloud Apps portal [8].
13. **Create Session Policy** for file upload inspection:
    - Activity filter: Upload file
    - Inspection method: Data Classification Service
    - Sensitivity label filter: Confidential, Highly Confidential
    - Action: Block with user notification [8][26]
14. **Create Session Policy** for file download protection:
    - Control file download (with inspection)
    - Protect action: Apply sensitivity label to downloaded file [8][26]

15. **Note limitation**: This covers only web browser–based access. Supplement with Endpoint DLP for native app paths [26].

Phase 4: Purview Compliance API Connector (Weeks 4–6)

16. **Enable pay-as-you-go billing** on Azure subscription linked to Purview tenant [4][7].
17. **Request Compliance API key scope** from OpenAI via `support@openai.com` with key name, last four digits, and scope (read) [7].
18. **Store API key in Azure Key Vault**; configure Key Vault integration for Purview Data Map [4][47].
19. **Assign Graph/Purview API roles** to managed identity via Cloud Shell: `Purview.ProcessConversationMessages.All` and `User.Read.All` [4][47].
20. **Add ChatGPT Enterprise data source to Purview Data Map** with OpenAI-provided workspace ID; configure and run initial full scan [47].
21. **Create collection policy** via DSPM for AI one-click ("Capture interactions for enterprise AI apps") [4][2].
22. **Configure Communication Compliance policy** for ChatGPT Enterprise: select Enterprise AI Apps channel, enable detection of sensitive information sharing, hate/discrimination, corporate sabotage, credential sharing [3][4][21].
23. **Configure Insider Risk Management** "Risky AI usage" policy template; enable Adaptive Protection to escalate DLP enforcement for elevated-risk users [1][4][48].
24. **Configure retention policies**: Select "Enterprise AI apps" location; set retention period per regulatory requirements [1].

Phase 5: Enforcement and Monitoring (Ongoing)

25. **Escalate Endpoint DLP from audit to block mode** (with override option initially) after policy tuning [27][39].
26. **Connect audit logs to SIEM** (Sentinel, Splunk, Datadog) using OpenAI-provided PowerShell/bash download scripts [18][42].
27. **Deploy DLP alert integration with Insider Risk Management** (now GA): DLP alerts feed into IRM indicators, eliminating the need to switch between portals

[48].

28. **Implement Adaptive Protection**: Configure DLP action tiers based on IRM risk level—minor risk gets warning, elevated risk gets hard block [7][48].
29. **Establish monthly review cadence**: Review Activity Explorer reports, DSPM for AI metrics, false positive rates, and policy coverage for new AI tools [53].

Part V: Licensing Summary

Control Layer	Minimum Licensing	Approximate Cost (list)	Notes
Purview Compliance API Connector	M365 E5 or Purview Suite add-on + Azure subscription	E5: $60/user/month; Purview Suite: $12/user/month [39][52]	Plus PAYG metered charges per prompt/response [7][39]
Endpoint DLP (paste/upload block)	M365 E5, E5 Compliance, or Purview Suite add-on	E5: $60; Purview Suite: $12/user/month [39][52]	E3 base covers basic sensitivity labels only [15]
Browser DLP (Edge for Business)	M365 E3 + Intune + Purview Suite	Intune P1: ~$8/user/month	Edge native; Chrome/Firefox require extension [15][49]
Communication Compliance	M365 E5, E5 Compliance, F5 Compliance	Included in E5; Purview Suite add-on	PAYG required for non-M365 data detection [21]
Insider Risk Management	M365 E5 or E5 Compliance	Included in E5	Risky prompt detection on third-party AI apps is metered PAYG (GA April 2026) [7]
eDiscovery (Standard)	M365 E3	Included	
eDiscovery (Premium)	M365 E5	Included in E5	
DSPM for AI	M365 E5 or Purview Suite	$12/user/month (Purview Suite) [39]	Business Premium users: $10/user/month [39][52]

Control Layer	Minimum Licensing	Approximate Cost (list)	Notes
Defender for Cloud Apps (MCAS)	M365 E5, EMS E5, or standalone MDCA	Included in E5 bundles [29][41]	Entra ID P1 additionally required for Conditional Access App Control [26]
Netskope (AI Gateway)	Separate SSE contract	Enterprise pricing (negotiated)	Netskope One Advanced SSE required [9]
Cyberhaven	Separate contract	Enterprise pricing	Available now for ChatGPT Enterprise [36]
Nightfall AI	Separate contract	Per-seat pricing	Browser plugin + endpoint agent [24]
Zscaler (ZIA Data Protection)	Separate ZIA contract	Enterprise pricing	WebSocket inspection requires Support enablement [52]

Part VI: Trade-Offs and Decision Framework

Scenario-Based Recommendations

Scenario A: Microsoft-First Organization with E5 Licensing, M365 Copilot as Primary AI Tool

Recommendation: Extend Purview. Prioritize M365 Copilot as the primary AI assistant (native DLP, full governance). For ChatGPT Enterprise usage, deploy Endpoint DLP + Browser DLP + Compliance API Connector. The Purview Suite add-on ($12/user/month) is likely already justified for other compliance use cases [35].

Trade-off: Desktop ChatGPT native app and non-browser paths remain outside Endpoint DLP's reach for typed prompts (only paste and upload are covered outside Edge) [52]. Accept this residual risk or supplement with Cyberhaven/Nightfall.

Scenario B: Regulated Industry (Financial Services, Healthcare, Legal) Requiring Preventive Controls

Recommendation: Layered enforcement is non-negotiable. Deploy Endpoint DLP + Browser DLP + MCAS session controls for web access. Supplement with a third-party tool (Netskope, Cyberhaven, or dope.security) for broader coverage including native apps and personal account differentiation. Use Purview Compliance API Connector for eDiscovery and Communication Compliance audit trail [16][35].

Trade-off: Multiple vendor complexity and cost. Netskope and Zscaler introduce network backhaul latency [13][50]. MCAS session controls have the 50 MB file and 30 MB inspection limits documented above [38][26].

Scenario C: Organization Using ChatGPT Enterprise Primarily via Browser on Intune-Managed Windows Devices

Recommendation: This is the optimal configuration for Microsoft-native controls. Deploy Endpoint DLP with paste/upload blocking, configure Edge for Business with Inline web traffic policy, add MCAS session controls for file inspection, and connect the Compliance API Connector for governance. This achieves near-complete coverage for the web-browser use case at minimal additional tool cost [49][15][7].

Trade-off: Any use of the native ChatGPT desktop app or API-based integrations falls outside this coverage. Monitor Shadow IT for these paths.

Scenario D: Heterogeneous AI Environment (ChatGPT, Claude, Gemini, Internal Models)

Recommendation: Purview alone is insufficient across the portfolio. Purview's unmanaged AI app browser protection covers only 19 AI apps [35]; a third-party tool with 1,400+ AI app coverage (such as dope.security or Netskope AI Gateway) is warranted for breadth. Use Purview for M365 Copilot enforcement and governance; use AI-native DLP for the heterogeneous surface [35].

Known Limitations Summary (Consolidated)

Limitation	Affected Controls	Mitigation
No real-time DLP in ChatGPT Enterprise natively	Purview Compliance API Connector	Endpoint DLP + Browser DLP + MCAS
24-hour ingestion delay for Purview connector	Purview Compliance API Connector	Supplement with SIEM for near-real-time alerts

Limitation	Affected Controls	Mitigation
Desktop AI native apps bypass browser/session controls	Browser DLP, MCAS Session Controls	Endpoint DLP with clipboard inspection; Cyberhaven agent
MCAS session controls cover only web browsers	MCAS Session Controls	Endpoint DLP for native app paths
50 MB file size limit for MCAS session policies	MCAS Session Controls	Tenant-level block for oversized files
Typed prompts covered in Edge only (not other browsers) for inline DLP	Endpoint DLP	Unallowed Browsers list + browser extension for Chrome/Firefox
macOS lags behind Windows for Endpoint DLP	Endpoint DLP	Supplement with third-party endpoint agent on macOS
BYOD and unmanaged devices outside Endpoint DLP scope	Endpoint DLP	MCAS session controls (no device agent required)
B2B guest users not covered by session policies or browser DLP	MCAS, Browser DLP	Separate policy design for guest access paths
Personal vs. enterprise account differentiation on shared URLs	Purview Endpoint DLP	Netskope, Zscaler, or dope.security for tenant differentiation
DLP cannot scan files uploaded directly into Copilot prompts	M365 Copilot DLP	Only typed text is evaluated; sensitivity label encryption provides indirect protection
Purview browser/network data security does not apply to B2B guests	Purview Browser/Network DLP	Acknowledge as known gap
EU compliance gaps: Schrems II TIA, DPIA lineage, EU AI Act Annex III	ChatGPT Enterprise governance	Engage legal counsel; consider supplementary measures (EKM, pseudonymization) [42]
Compliance API key requires email request to OpenAI (no self-service)	Purview Connector, third-party tools	Build into onboarding SLA expectations

Limitation	Affected Controls	Mitigation
File policies in MCAS retire January 6, 2027	MCAS File Policies	Migrate to Purview DLP or auto-labeling before deadline [28]

Part VII: Forward-Looking Considerations

Several capabilities are in active development that will materially change this architecture:

- **Purview + Microsoft Entra Global Secure Access integration**: GA announced for September 2026, enabling network-layer DLP for text, files, and AI interactions intercepted at the network layer—including protection of generative AI platforms [20]. This will reduce dependence on MCAS session controls for browser-based AI traffic.
- **Inline prompt protection for Microsoft Foundry apps** via Purview `processContent` API: GA expected June 2026 [7]. If OpenAI adopts the same `processContent` integration, it would enable true inline, preventive DLP for ChatGPT Enterprise at the platform level—the most significant potential architectural change on the horizon.
- **Purview-Island Enterprise Browser integration**: Public preview since late March 2026; GA expected June 2027. Extends Purview DLP to Island-managed browser workflows including typed/pasted input and file transfers at the presentation layer [30].
- **Anthropic Claude Enterprise** will gain full Purview support (classification, IRM, Communication Compliance, eDiscovery, Data Lifecycle Management, Compliance Manager) as of August 2026 [20], reducing the gap between Claude and ChatGPT Enterprise in the Purview ecosystem.
- **MCAS file policies retirement** (January 6, 2027): Organizations must migrate to Purview DLP or auto-labeling policies [28].
- **MCP (Model Context Protocol) security**: AI agent-to-enterprise-data connections via MCP are not currently addressed by any published Purview DLP documentation and represent an emerging gap as MCP adoption accelerates [53]. Third-party tools (Strac, Cyberhaven) have begun shipping MCP DLP integrations.

Conclusion and Final Recommendations

1. **Do not treat the Purview Compliance API Connector as DLP**. It is an essential governance tool for audit, eDiscovery, retention, Communication Compliance, and Insider Risk Management—but it prevents nothing [1][7][35]. Set stakeholder expectations accordingly before deployment.
2. **The closest approximation to true DLP for ChatGPT Enterprise within the Microsoft stack** is a combination of Endpoint DLP (paste/upload blocking on managed devices) + Browser DLP in Edge for Business (typed prompt blocking) + MCAS session controls (file upload inspection for web sessions). This combination achieves meaningful prevention but carries the documented limitations above [7][15][26][49].
3. **Microsoft 365 Copilot has fundamentally stronger native DLP coverage** than ChatGPT Enterprise. For organizations where GPT capabilities are primarily needed for internal M365 data tasks, Copilot is the more governable choice. For organizations requiring ChatGPT Enterprise specifically (third-party model, OpenAI capabilities, custom GPTs), accept the governance gap and mitigate through layered controls [5][7][12].
4. **For organizations with strict preventive requirements**, supplement the Microsoft stack with a third-party AI-native DLP tool. Cyberhaven offers the deepest data lineage and ChatGPT Enterprise Compliance API integration. Netskope AI Gateway provides inline inspection with personal/corporate account differentiation. Nightfall provides rapid browser deployment with ML-native detectors. dope.security provides on-device inspection without backhaul latency [36][46][50][13].
5. **Deploy in phases**: Start with governance (Compliance API Connector + Communication Compliance + IRM) while endpoint and browser controls are configured. Deploy Endpoint DLP in audit mode before block. Expand to MCAS session controls. Add third-party tooling if coverage gaps remain after assessing the Microsoft-native stack against actual risk profile.
6. **Prioritize the endpoint plane over the browser plane** for highest-risk users: Endpoint DLP takes precedence over MCAS session policies when both are

configured for the same context [26], and endpoint-level controls capture clipboard paste that browser session controls may miss.

7. **Monitor the Purview roadmap closely**: The potential adoption of the `processContent` inline API by OpenAI would fundamentally change Option 1 from post-hoc governance to real-time prevention—making the Purview connector the most important future development to track [7].

Note: Capability status, licensing prices, and product availability are as of mid-2026 based on available evidence. Microsoft Purview and OpenAI product capabilities are evolving rapidly; verify current status of preview features before deployment decisions.

Source links

1. https://learn.microsoft.com/en-us/purview/ai-chatgpt-enterprise
2. https://learn.microsoft.com/en-us/purview/archive-chatgpt-interactions
3. https://techcommunity.microsoft.com/blog/microsoft-security-blog/microsoft-purview-integrates-with-chatgpt-enterprise-compliance-api-to-support-c/4192868
4. https://techcommunity.microsoft.com/blog/microsoft-security-blog/unlocking-the-power-of-microsoft-purview-for-chatgpt-enterprise/4371239
5. https://learn.microsoft.com/en-us/purview/dlp-microsoft365-copilot-location-learn-about
6. https://help.openai.com/en/articles/9261474-openai-compliance-platform-for-enterprise-and-edu-customers
7. https://calebamcdowell.substack.com/p/extending-purview-into-enterprise
8. https://learn.microsoft.com/en-us/defender-cloud-apps/session-policy-aad
9. https://blog.admindroid.com/how-to-prevent-users-from-uploading-sensitive-data-to-chatgpt/
10. https://openai.com/index/new-tools-for-chatgpt-enterprise/

11. https://accuroai.co/blog/microsoft-copilot-vs-chatgpt-enterprise-data-leakage
12. https://www.tekstack.com/2025/11/13/copilot-vs-chatgpt-why-enterprise-data-protection-matters/
13. https://dope.security/post/chatgpt-dlp
14. https://thalpius.com/2026/05/31/blocking-generative-ai-with-microsoft-defender-for-cloud-apps-and-microsoft-defender-for-endpoint/
15. https://blog.admin365.ai/prevent-users-from-uploading-sensitive-data-to-chatgpt-using-microsoft-purview/
16. https://www.proarch.com/blog/microsoft-purview-chatgpt-ai-data-protection
17. https://intuitionlabs.ai/articles/chatgpt-enterprise-admin-controls-security
18. https://www.harmonic.security/resources/securing-chatgpt-enterprise-guide
19. https://chatgpt.com/business/enterprise/
20. https://learn.microsoft.com/en-us/purview/whats-new
21. https://blog.admindroid.com/find-ai-interactions-with-communication-compliance-policy-in-microsoft-purview/
22. https://learn.microsoft.com/en-us/answers/questions/2154287/blocking-file-uploads-to-generative-ai-websites-us
23. https://kcpdynamics.com/en/dlp-in-copilot-secure-configuration-microsoft-365/
24. https://www.nightfall.ai/lp/chatgpt-dlp
25. https://dope.security/post/top-ai-dlp-tools-chatgpt-claude-2026
26. https://www.tiagoscarvalho.com/security-compliance/dlp-vs-session-controls-2026
27. https://nikkichapple.com/microsoft-purview-dlp-copilot-genai-setup-guide/
28. https://learn.microsoft.com/en-us/defender-cloud-apps/content-inspection
29. https://www.epcgroup.net/microsoft-defender-cloud-apps-casb-saas-security-2026

30. https://mc.merill.net/message/MC1249424
31. https://officegarageitpro.medium.com/data-security-for-agents-and-3rd-party-ai-in-microsoft-purview-1aaaeb54c310
32. https://aona.ai/blog/chatgpt-enterprise-security-guide/
33. https://www.strac.io/blog/microsoft-endpoint-data-loss-prevention
34. https://docs.paloaltonetworks.com/enterprise-dlp/administration/configure-enterprise-dlp/enterprise-dlp-and-ai-apps/how-enterprise-dlp-safeguards-against-chatgpt-data-leakage
35. https://accuroai.co/blog/purview-vs-ai-native-dlp
36. https://www.cyberhaven.com/press-releases/cyberhaven-extends-ai-security-chatgpt-enterprise-claude
37. https://dlptest.com/cyberhaven-extends-data-security-platform-chatgpt-enterprise-claude/
38. https://learn.microsoft.com/en-us/defender-cloud-apps/caac-known-issues
39. https://o365hq.com/blog/microsoft-purview-dspm-for-ai-what-it-shows-licensing-and-setup
40. https://mc.merill.net/message/RM401125
41. https://blog.ciaops.com/2026/06/04/defender-for-cloud-apps-session-policy/
42. https://www.teamazing.com/blog/chatgpt-enterprise-eu-review-2026/
43. https://learn.microsoft.com/en-us/purview/dspm-for-ai
44. https://learn.microsoft.com/en-us/purview/ai-microsoft-purview
45. https://info.adaptivedge.com/blog/restrict-pasting-sensitive-information-to-browser-sites-with-purview-endpoint-dlp
46. https://www.netskope.com/resources/demos-and-videos/netskope-chatgpt-enterprise-compliance-api
47. https://loaseguraria.com/2025/05/08/monitoring-chatgpt-enterprise-with-purview/

48. https://techcommunity.microsoft.com/blog/microsoft-security-blog/mitigating-insider-risks-in-the-age-of-ai-with-microsoft-purview-insider-risk-ma/4396057
49. https://mertefekanlikilic.com/how-to-manage-ai-chatbot-access-with-purview-dlp-and-intune/
50. https://dope.security/post/best-ai-dlp-software-chatgpt-claude-compared-2026
51. https://www.cyberhaven.com/blog/best-enterprise-dlp-tools-ai-data-risk
52. https://jazz.security/dlp-explained/dlp-basics/best-dlp-software/
53. https://www.strac.io/blog/ai-dlp