DEEPWATCH™
Always Watching. Always Protecting.

www.deepwatch.com

Strategy Guide:

Copilot Security Threats and Recommendations

Published: Mar 10, 2026

Analysis by:
Deepwatch Threat Intelligence

Executive Summary

This brief strategy guide outlines the critical security considerations for deploying Windows and Microsoft 365 Copilot in an enterprise environment as of January 2026. While Copilot offers productivity gains, it acts as a "permission amplifier," necessitating a Zero Trust approach to prevent data exposure and attacks.

The primary risk of Generative AI (GenAI) in Windows 11 and Microsoft 365, is not the AI model itself, but the amplification of existing configuration gaps. In 2026, organizations must move beyond basic "Shadow AI" blocking to active Data Security Posture Management (DSPM).

Primary Threat Vectors

1. Permission Amplification & Over-sharing

Copilot operates using the "on-behalf-of" model, meaning it has access to every file, email, and chat the user can technically access.

- **The Risk:** Most enterprises have "permission sprawl," where users have access to sensitive HR or financial folders they never actually open. Copilot makes this "dark data" discoverable via a simple natural language query (e.g., "Summarize the salary spreadsheets I have access to").
- **Impact:** Massive internal data leakage and accidental insider threats.

2. Prompt Injection & "Reprompt" Attacks

A significant evolution in 2025-2026 has been the rise of Indirect Prompt Injection.

- **The Risk:** An attacker sends a malicious email or hides invisible text on a website. When Copilot "reads" that content to summarize it for a user, the hidden instructions take over the session.
- **2026 Update:** Recent "Reprompt" vulnerabilities allow attackers to bypass standard filters by chaining multiple requests, potentially siphoning data to external servers even after a chat is closed.

3. Data Residency and Compliance Drift

- **The Risk:** While Microsoft guarantees that enterprise data is not used to train foundational models, "Web Grounding" (using Bing to improve answers) can inadvertently leak context to the public web if not strictly configured.
- **Compliance:** Failure to align Copilot with the EU AI Act or GDPR (regarding the "Right to be Forgotten" within AI-generated summaries) poses significant legal risks.

Strategic Recommendations

To mitigate these risks, organizations should follow a three-phased "Secure-by-Design" rollout:

Phase 1: Foundation (Data Hygiene)

- **Enforce Least Privilege:** Use tools like Microsoft Purview to identify over-shared files and "just-enough-access" (JEA) policies.
- **Enable "Restricted Search Mode" (Tenant-Wide):** If you are worried about widespread over-sharing, you can limit Copilot to only search a "curated list" of allowed sites. O PowerShell can be used to limit Copilot's reach to only explicitly allowed sites O Set-SPOTenant -RestrictedSearchMode $true
- **Sensitivity Labeling:** Automate the labeling of sensitive data. Copilot respects these labels and will refuse to summarize or output content marked as "Confidential" if configured correctly. You can use PowerShell to bulk-apply "Confidential" labels to known sensitive folders.
- **Disable Web Search (Initially):** In highly regulated sectors, disable Bing integration within Copilot settings to ensure a closed-loop environment.

Phase 2: Technical Guardrails

As of **Q1 2026**, Microsoft has streamlined the administration of Copilot, but because it is woven into so many different layers, there is no single "Off" switch.

To disable Copilot features for a phased rollout, you must address the **three primary control planes**: the Microsoft 365 Admin Center, the Power Platform Admin Center, and Device Management (Intune/GPO).

1. Microsoft 365 Apps (Word, Excel, PPT, Outlook, Teams)

If your users have licenses but you aren't ready for them to see the AI buttons in their apps:

- **The License Method:** The most effective way to "disable" it is to **unassign the Microsoft 365 Copilot license** from the user in the **Microsoft 365 Admin Center** (Users > Active users).
- **The Policy Method (Cloud Policy):** If you want to keep the licenses assigned but hide the features:
    - 1. Go to the Microsoft 365 Cloud Policy Service.
    - 2. Create a policy configuration.
    - 3. Search for and disable: **"Enable Copilot"**. This will remove the Copilot ribbon icons across Word, Excel, PowerPoint, and OneNote.
- **Outlook Specifics:** In the **Exchange Admin Center**, you can manage the "Copilot in Outlook" app/add-in to prevent it from appearing in the mail client.

2. Windows 11 & Microsoft Edge

These are handled via endpoint management. In early 2026, a new "clean" removal policy was introduced.

Windows 11 (Intune):

- Navigate to *Devices > Configuration > Create > New Policy.*
- **Platform:** Windows 10 and later | **Profile:** Settings Catalog.
- Search for **"Turn off Copilot in Windows"** (under the **Windows AI** category) and set it to **Enabled**.

New for 2026 (GPO):

A specific policy called **"Remove Microsoft Copilot App"** was added to Windows 11 builds in January 2026.

- **Path:** *User Configuration > Administrative Templates > Windows AI > Remove Microsoft Copilot App.*

Microsoft Edge:

- Use the Edge Management service or Intune Settings Catalog.
- Search for **"Allow Copilot in Microsoft Edge"** and set it to **Disabled**.

3. Data & Business Apps (Fabric, Power Platform, Dynamics)

These require visits to their specific specialized admin portals.

Microsoft Fabric:

1. Go to the **Fabric Admin Portal**.
2. Navigate to **Tenant Settings**.
3. Find **"Users can use Copilot and other features powered by Azure OpenAI"** and toggle it to **Disabled**.

Power Apps & Power Automate:

1. Open the **Power Platform Admin Center**.
2. Go to Settings > Tenant Settings.
3. Select **"Copilot in Power Apps (preview)"** or **"Copilot in Power Automate"** and toggle them **Off**.

Note: You can also do this at the **Environment** level under Settings > Product > Features.

4. Security & Governance (Defender, Purview, Entra, Intune)

For the security suite, Copilot usually operates as an "embedded" assistant.

- **Security Copilot Portal:** Access to the full standalone Security Copilot is governed by **Entra ID Security Groups**. If a user isn't in the "Security Copilot Users" group, they cannot access the interface.
- **Intune/Defender Integration:** You can disable the "embedded" insights by going to the **Microsoft Security Copilot portal** and disabling the specific **Plug-ins** for Intune, Defender, or Entra. This removes the "Summarize this policy" or "Analyze this script" buttons from those admin centers.

Summary Checklist for Admins

Surface	Admin Portal	Primary Setting/Policy
M365 Apps	config.office.com	"Enable Copilot" (Set to Disabled)
Windows 11	Intune / GPO	"Turn off Copilot in Windows"
Edge Browser	Edge Management	"Allow Copilot in Microsoft Edge"
Fabric / Power BI	Fabric Admin Portal	"Users can use Copilot..." (Tenant level)
Power Platform	PP Admin Center	"Copilot (preview)" (Tenant/Env level)
Security Suite	Security Copilot Portal	Manage/Disable Source Plug-ins

Note: *PowerShell can be a useful tool to bulk-assign/unassign Copilot licenses or check which users currently have them active.*

Phase 3: Continuous Governance

Governance is an ongoing process, not a one-time configuration. A crucial consideration in 2026 is "model drift," where AI models' performance or safety alignment changes as they are updated or exposed to new data. A successful and secure AI implementation hinges on several critical components: user education, comprehensive auditing, diligent logging, and proactive red teaming of your AI implementation.

Control Category	Recommendation
Audit Logging	Review Purview Audit logs weekly for anomalous AI queries (e.g., a junior dev asking about executive bonuses). Audit Regularly: Review Search-UnifiedAuditLog weekly to identify users attempting to bypass AI guardrails via "jailbreak" prompts.
User Education	Train staff on "Prompt Safety." Teach them to never paste AI outputs into external public-facing tools.
Red Teaming	Conduct "AI Red Teaming" exercises to test if your internal controls can stop a simulated prompt injection attack.

Consider Adding These Items to Your 2026 Security Checklist

- Tenant Isolation Verified: Ensure no cross-tenant data sync is enabled.
- DLP Policies Updated: Update Data Loss Prevention (DLP) rules to specifically recognize and block AI-generated summaries containing PII.
- Shadow AI Discovery: Use Defender for Cloud Apps to identify employees using unapproved, non-enterprise AI versions.
- Zero Trust for AI: Treat every Copilot response as "untrusted" until verified by a human (human-in-the-loop).