

Cybersecurity Assessment of Microsoft Co-Pilot

Cybersecurity Risk Assessment Report

Table of Contents

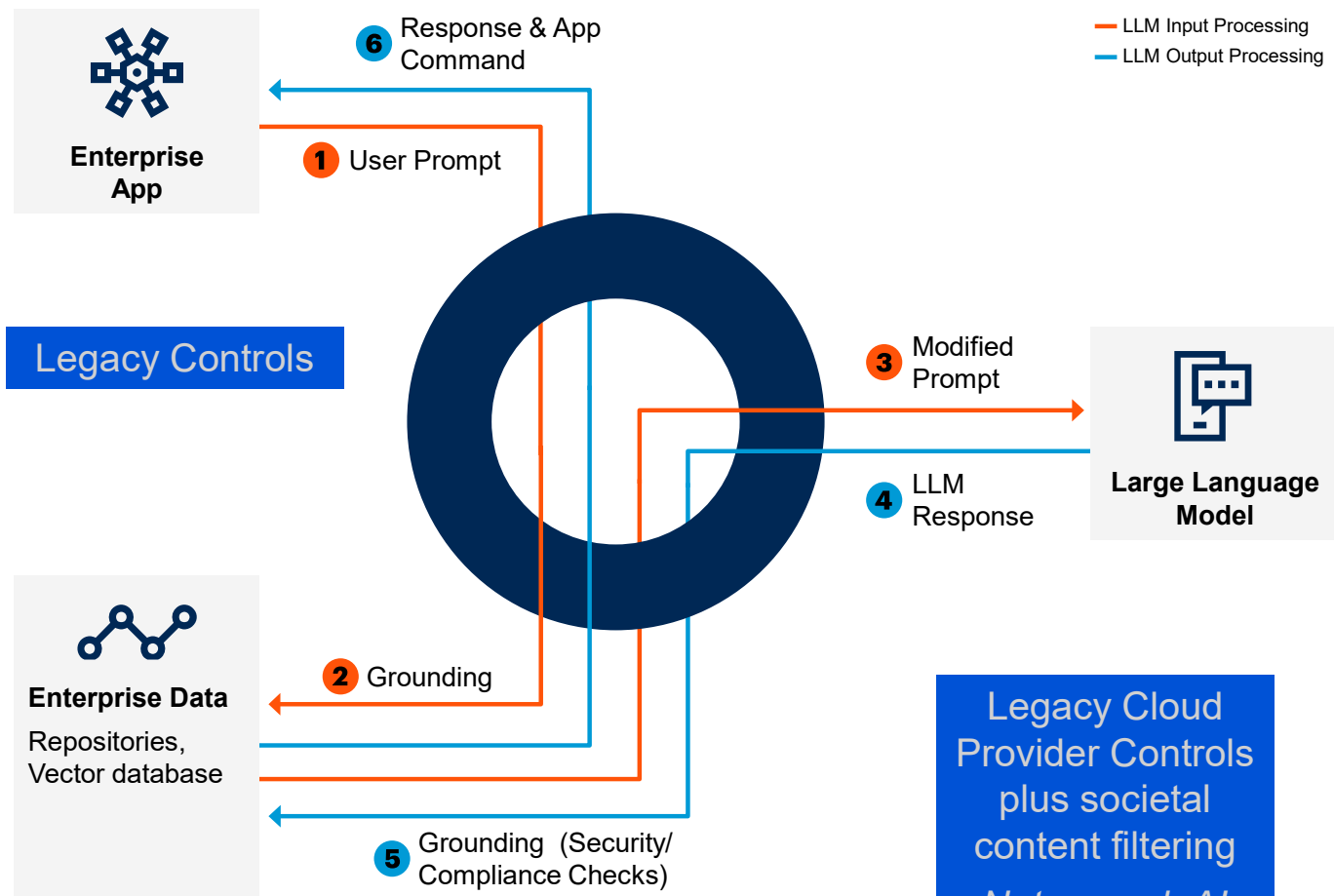
- Scope Summary
- Assessment Approach
- Generic Large Language Model & Microsoft Co-Pilot Architecture
- Summary of Findings
- Recommendations

Assessment Approach

The Cybersecurity Assessment was conducted by understanding the Microsoft 365 suite for Co-Pilot, M365 Chat and interviews with Co-Pilot Core Team and Microsoft. Reviewing Gartner reports and overall understanding of risks related to Generative AI and Large Language Models. Reviewed AI guidance from Canadian Cyber Security Center, Organization's AI Policy, FBI and US Congress Primers on AI and Microsoft Links on use of Responsible AI.

Generic Large Language Model-Enterprise Use Case example

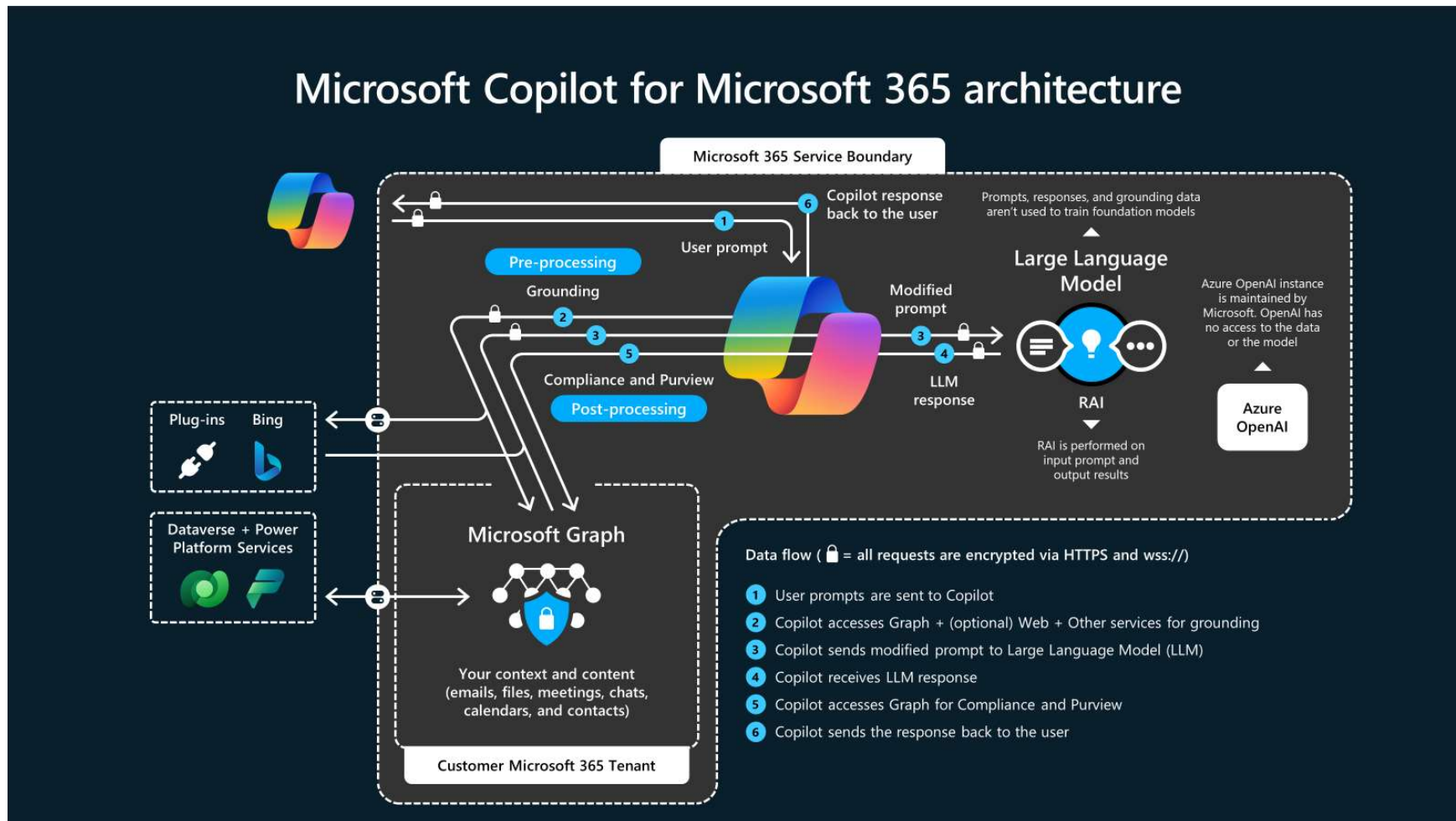
Legacy Controls are not enough



Source: Gartner, using information from Microsoft
793862_C

Legacy Cloud Provider Controls plus societal content filtering
Not enough AI-specific controls

Microsoft Co-Pilot Architecture



Summary of Findings

Summary of Issues and Observations

Risk Assessment Conclusion

Based on the risk assessment performed for Microsoft Co-Pilot, following findings are noted below

Findings Summary					
Total Issues Identified	Very High	High	Medium	Low	Very Low
	0	1	3	1	0

Co-Pilot Technology Risk Assessment

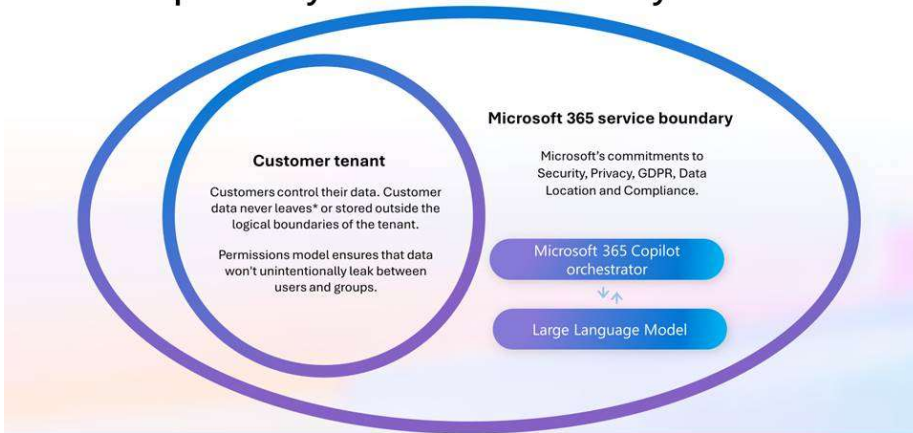
Summary: Cyber Recommendations for Internet Access of Co-Pilot

- **Co-Pilot as a technology has well established technical security controls under Microsoft purview as a SaaS platform including data encryption at rest, in motion and access controls on a need-to-know basis for end users.**
- **There is minimal risk in turning “ON” internet access to Co-Pilot as the data transfer happens via encrypted channels and there is no storage of organization specific data on Microsoft Bing internet services used in this case. (see slide 6, Co-Pilot Architecture for details)**

Microsoft Safeguards

- Commercial data protection means chat data isn't saved, Microsoft has no eyes-on access, and your data isn't used to train the underlying large language models. Organizations can learn more about what commercial data protection provides here: [Bing Chat Enterprise Privacy and Protections | Microsoft Learn](#)

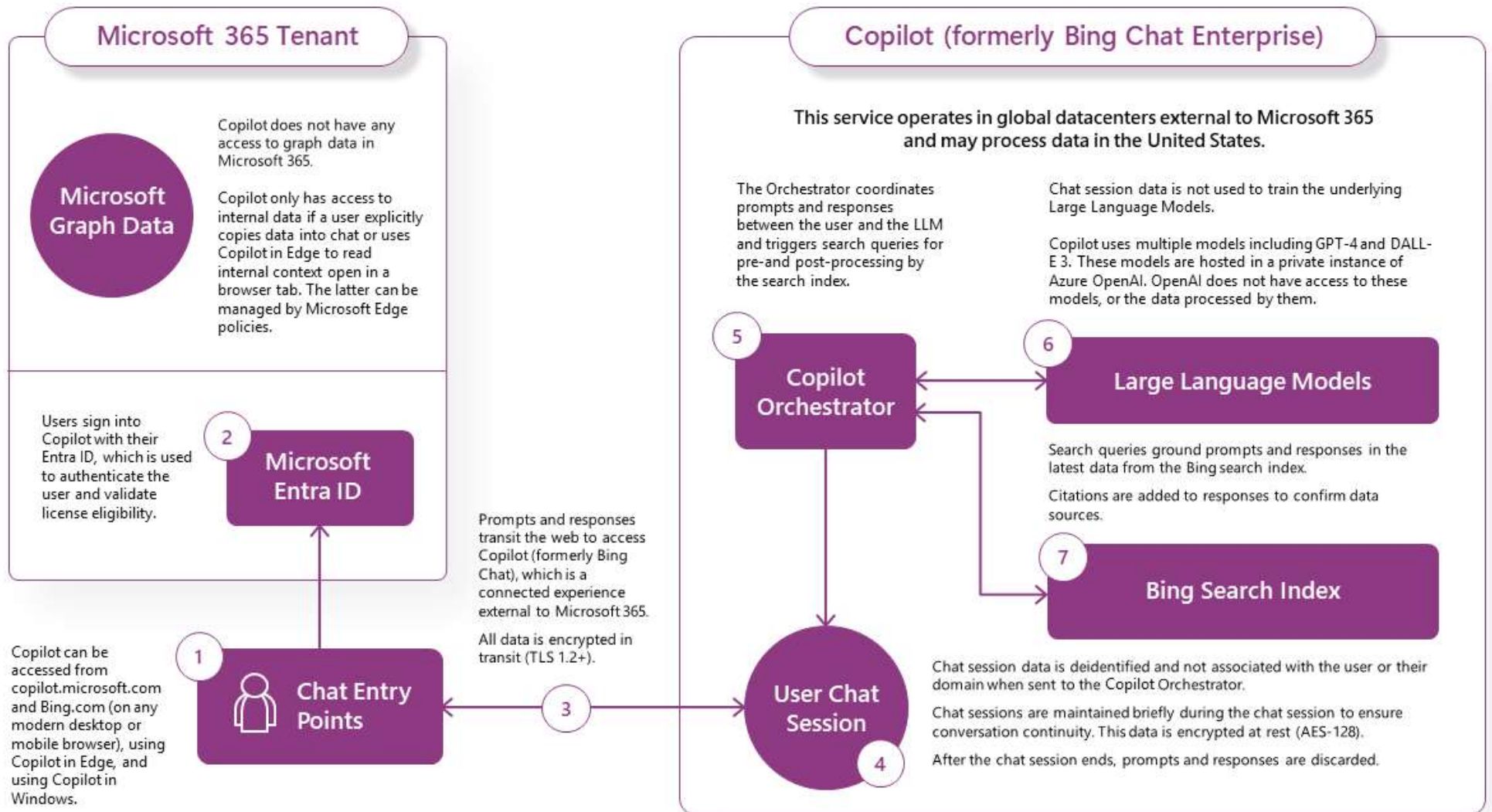
Microsoft protects your data and enables you to control it



*unless a user searches the web, which can be turned off, or customer consent is given

📌 Important

- Microsoft Copilot for Microsoft 365 is compliant with our existing privacy, security, and compliance commitments to Microsoft 365 commercial customers, including the General Data Protection Regulation (GDPR) and European Union (EU) Data Boundary.
- Prompts, responses, and data accessed through Microsoft Graph aren't used to train foundation LLMs, including those used by Microsoft Copilot for Microsoft 365.



Data Residency within Microsoft User Tenant

- **When my users are signed in with Entra ID, does data put into Copilot stay within my Microsoft 365 tenant boundary?**
- Copilot is a connected service where Microsoft is the data controller. Entra ID users' prompts leave your organization's Microsoft 365 tenant boundary to reach the Copilot service. However, this data is encrypted in transit, and Microsoft doesn't retain this data beyond a short caching period for runtime purposes. After the browser is closed, the chat topic is reset, or the session times out, Microsoft discards all prompts and responses.
- To provide chat responses, Copilot uses global data centers for processing and may process data in the United States. Optional, Bing-backed connected experiences don't fall under Microsoft's EU Data Boundary (EUDB) commitment. They also don't fall under the Data Protection Addendum, which is specifically reserved for services where Microsoft is just the data processor.
- As a reminder, Copilot has no access to organizational data inside your tenant boundary, and chat conversations aren't saved or used to train the underlying models.
- Organizations with strict requirements that data must remain inside tenant or geographic boundaries should instead consider Copilot for Microsoft 365 or Azure Open AI to provide generative AI services. Copilot with commercial data protection is intended as a more secure alternative for organizations over consumer generative AI services..


Co-Pilot Information Governance Risks


Summary of Information Governance Risks

- **Risk of the use of Co-Pilot mostly focusses on the responsible use of the technology and the information sources feeding co-pilot which are oversharing, inaccuracy, information sprawl and mistakes (confidently made statements via hallucination on wrong data). Microsoft suggest their grounding mechanism reduces the hallucination problem.**
- **Poor information management undermines the value of Copilot for Microsoft 365 and other generative AI tools. It also increases risks of oversharing, misinformation and data loss. To mitigate risk, application technical professionals must establish effective information governance in Microsoft 365.**

Key Findings

- The value of Copilot for Microsoft 365 is highly dependent on the quality of information stored in the organization's Microsoft 365 environment and the controls in place to protect it.
- Information governance in Microsoft 365 remains a significant challenge. In Gartner's 2023 Microsoft 365 survey, almost 60% of respondents stated that oversharing, data loss and content sprawl were among the biggest risks to their organization's Microsoft 365 environment. Deploying Copilot for Microsoft 365 without first taking steps to address these risks, will increase them.
- While Copilot for Microsoft 365 respects user permissions, if content has been overshared, Copilot's response may contain information the user should not have access to. (*Privacy risk*)
- Copilot for Microsoft 365 makes the creation of complex objects such as SharePoint sites, Power Apps and Power Automate flows far easier. However, if left unchecked, this could result in a new era of AI-generated sprawl that organizations will be ill-equipped to manage.
- Co-Pilots prompts don't currently have content filtering offered by Microsoft which means prompts can be created to share and consume sensitive information not following information classification policies.

 Very high risk

 High risk

 Medium risk

 Low risk

 Very low risk

Recommendations

- Establish a clear information architecture, permission, and retention model for Microsoft 365 to ensure that content stored in OneDrive, Teams and SharePoint has appropriate access rights and life cycle rules.
- Classify and protect sensitivity information by using Microsoft Information Protection and Data Loss Prevention capabilities. This reduces the risk of oversharing and accidental discovery. Before deploying Copilot, test your access controls.

(see Appendix for practical recommendations on Information Governance in M365 which can be enabled as quick wins)

- Assess third-party add-on products to determine if they are required to improve visibility and governance of information across Microsoft 365.
- Develop mandatory training to help users understand how best to store and share information in Microsoft 365. Integrate this with your employee onboarding program and existing compliance and data security training.
- To manage AI related information sprawl risks, Organization should invest in a good user awareness training.

Recommendations

- To handle and mitigate Information Governance risks and ensure proper alignment in the use of Co-Pilot it is recommended that a work stream be established within TIS and Business to handle these risks. Co-Pilot work is not intended to solve these risks in silo and need to be wholistically looked at for their remediation.
- Promote a “distrust and verify” approach for content generated by Copilot for Microsoft 365 and other generative AI technologies. Users should not rely on the authenticity and accuracy of this content and should always validate the information and its sources.
- Revise Organization’s AI Policy to be in line with good Information governance practices. With pervasiveness of AI it is recommended to have Annual Attestation is done by Organization’s users similar to IT Acceptable Use Policy and Cyber Security Policies.
- Look on the possibility of introducing a banner on the use of Co-Pilot for Responsible Use of AI.

Appendix

Indirect Prompt Injection: the new attack vector

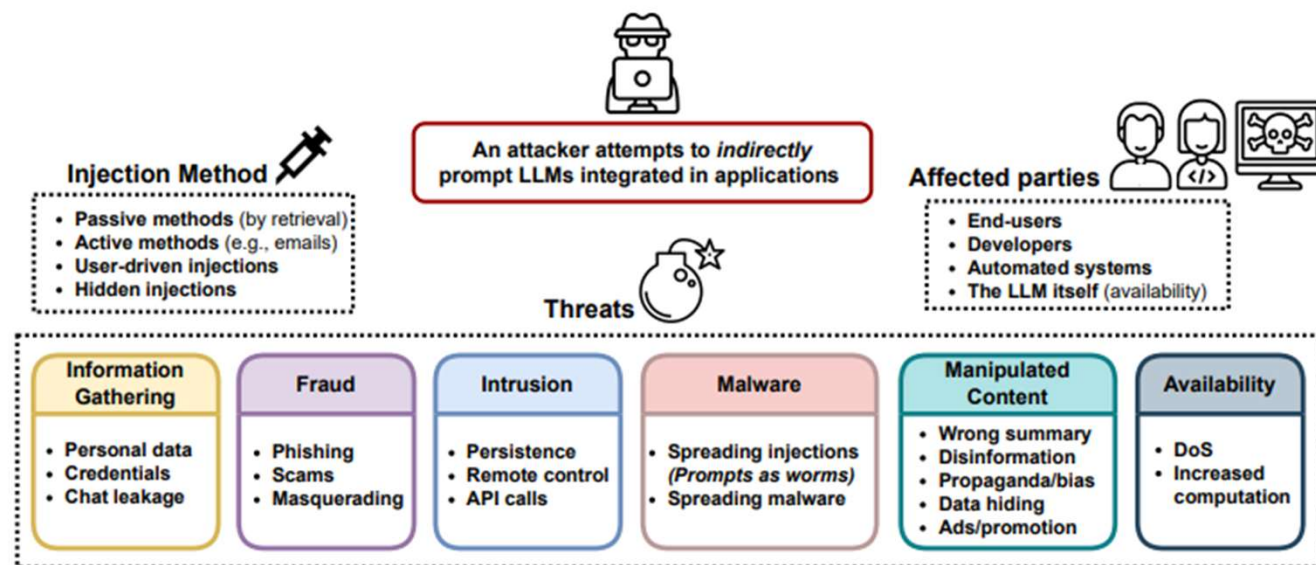


Figure 2: A high-level overview of new indirect prompt injection threats to LLM-integrated applications, how the prompts can be injected, and who can be targeted by these attacks.

Example: Injection Spreads Malware via Phishing

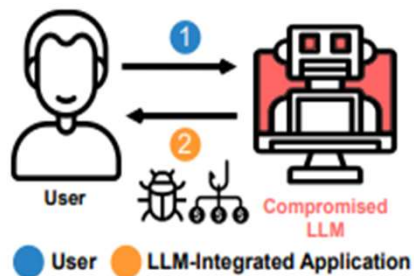


Figure 5: LLM-integrated applications can enable **fraud** and **malware** attacks. A user interacts with a compromised LLM ① that was prompted to distribute fraudulent or malicious links within its answers ②.

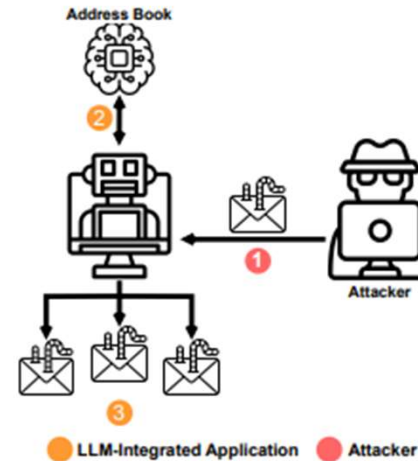






Figure 6: AI **malware**: the LLM-augmented email client receives an incoming email with a malicious payload ①, reads the user's address book ②, and forwards the message ③.

- Microsoft mentioned that Prompts including LLM is within Microsoft boundaries and is not externally exposed unlike Open AI models (see slide 12 on Microsoft safeguards)
- There is robust technical security controls to safeguard the LLM and thwart any Man-in-the-Middle attacks in Injecting malicious payloads in prompts and their return values.

Information Risks posed by Copilot for MS 365

Information Risks Posed by Copilot for Microsoft 365

 1	 2	 3	 4
Oversharing	Inaccuracy	Sprawl	Mistakes
Users may get responses based on information they should not have access to	Response could be based on stale or irrelevant content, spreading inaccuracies and misinformation	Copilot could lead to the rapid creation of content, containers and other M365 objects	Copilot could hallucinate and get things wrong

Source: Gartner
800669_C

Six Steps to Improve Microsoft 365 Information Governance

Six Steps to Improve Microsoft 365 Information Governance



Source: Gartner
800669_C

Gartner.

Gartner®

Figure 11: SharePoint Default Sharing Links

File and folder links

Choose the type of link that's selected by default when users share files and folders in SharePoint and OneDrive.

- Specific people (only the people the user specifies)
- Only people in your organization
- Anyone with the link

Choose the permission that's selected by default for sharing links.

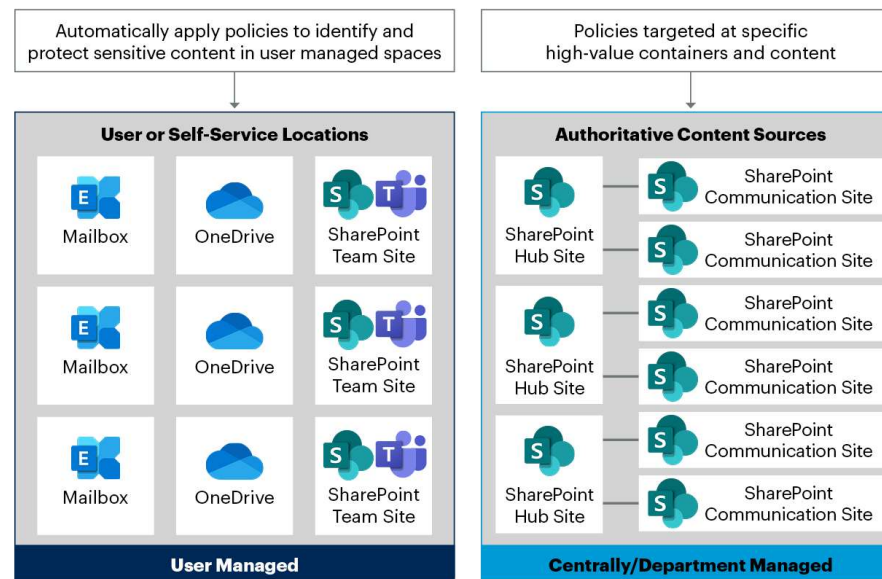
- View
- Edit

Gartner

Gartner®

Example of a Flat Architecture in Microsoft 365

Example of Flat Architecture in Microsoft 365



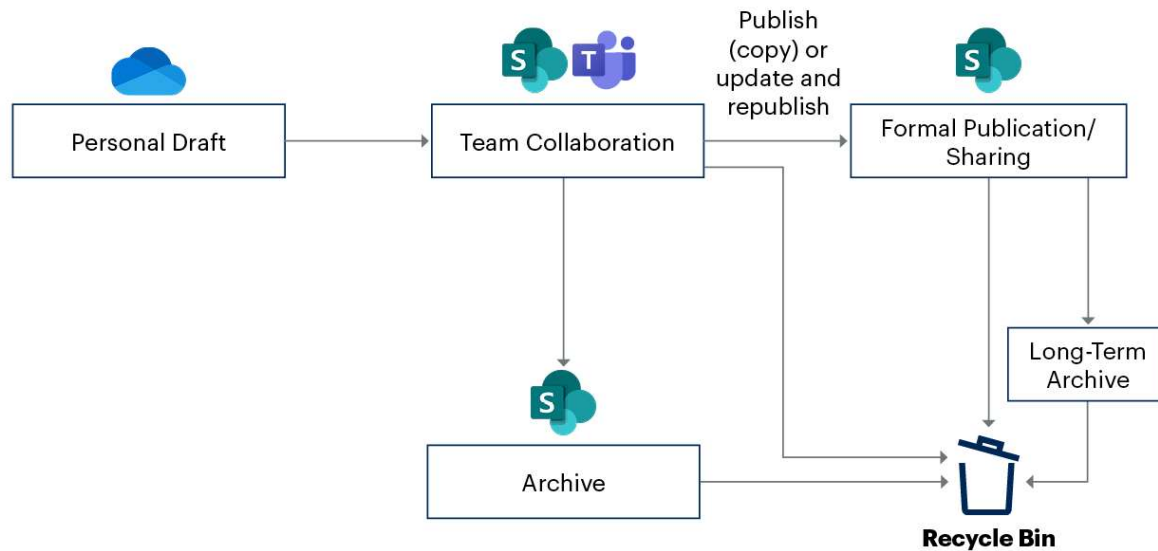
Source: Gartner
800669_C

Gartner

Gartner®

Example Content Life Cycle in Microsoft 365

Example Content Life Cycle in Microsoft 365



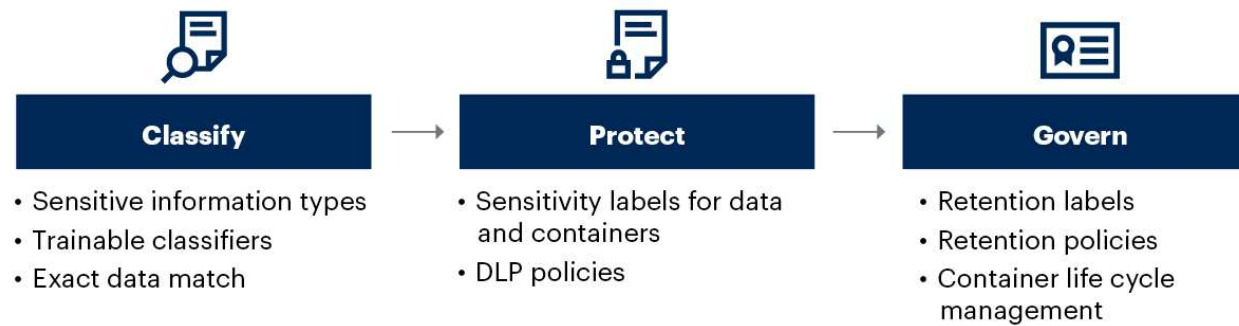
Source: Gartner
796809_C

Gartner

Gartner®

Classify, Protect and Govern Your Information

Classify, Protect and Govern Your Information



Source: Gartner
800669_C

Gartner

Gartner®

Example DLP Rule

We'll apply this policy to content that matches these conditions.

Content contains

Group name: Default Group operator: Any of these

Sensitive info types

Credit Card Number Medium confidence Instance count: 1 to Any

Actions

Restrict access or encrypt the content in Microsoft 365 locations

- Block users from receiving email or accessing shared SharePoint, OneDrive, and Teams files.
By default, users are blocked from sending Teams chats and channel messages that contain the type of content you're protecting. But you can choose who is blocked from receiving emails or accessing files shared from SharePoint, OneDrive, and Teams.
- Block everyone.
- Block only people outside your organization.

User notifications

Use notifications to inform your users and help educate them on the proper use of sensitive info. On

Annotations:

- Arrow pointing to "Credit Card Number": If content contains a credit card number
- Arrow pointing to "Block users from receiving email...": Access is blocked across M365 services
- Arrow pointing to "Use notifications to inform your users...": Owner is notified explaining why they cannot share the item.

Gartner

Copilot Prompt Example — Minimal Context

10/13 3:04 PM

Write Powershell script to get Microsoft 365 groups where AccessType = Public

M365 Chat 10/13 3:04 PM

Here is a PowerShell script that you can use to get Microsoft 365 groups where the AccessType is set to Public:

```
# Connect to Exchange Online
$UserCredential = Get-Credential
$Session = New-PSSession -ConfigurationName Microsoft.Exchange -ConnectionUri
https://outlook.office365.com/powershell-liveid/ -Credential $UserCredential -
Authentication Basic -AllowRedirection
Import-PSSession $Session

# Get all Microsoft 365 groups where AccessType is Public
$PublicGroups = Get-UnifiedGroup | Where-Object {$_.AccessType -eq "Public"}

# Display the results
$PublicGroups | Format-Table DisplayName, PrimarySmtpAddress

# Disconnect from Exchange Online
Remove-PSSession $Session
```

This script will prompt you to enter your credentials to connect to Exchange Online. Once connected, it will retrieve all Microsoft 365 groups where the AccessType is set to Public and display their display name and primary SMTP address in a table format. Finally, it will disconnect from Exchange Online.

1/30 AI-generated content may be incorrect

Gartner

Gartner®

Copilot Prompt Example — Expanded Context

10/13 3:09 PM

Write Powershell script to get Microsoft 365 groups where AccessType = Public which uses Modern Authentication

M365 Chat 10/13 3:09 PM

Here is a PowerShell script that you can use to get Microsoft 365 groups where the AccessType is set to Public using Modern Authentication:


```
# Connect to Exchange Online using Modern Authentication
Connect-ExchangeOnline



# Get all Microsoft 365 groups where AccessType is Public
$PublicGroups = Get-UnifiedGroup | Where-Object {$_.AccessType -eq "Public"}

# Display the results
$PublicGroups | Format-Table DisplayName, PrimarySmtpAddress

# Disconnect from Exchange Online
Disconnect-ExchangeOnline
```

This script uses the `Connect-ExchangeOnline` cmdlet to connect to Exchange Online using Modern Authentication. Once connected, it will retrieve all Microsoft 365 groups where the AccessType is set to Public and display their display name and primary SMTP address in a table format. Finally, it will disconnect from Exchange Online using the `Disconnect-ExchangeOnline` cmdlet.

 Copy

2/30 AI-generated content may be incorrect

Gartner.

Gartner®

References

- Generative Artificial Intelligence: [Generative artificial intelligence \(AI\) - ITSAP.00.041 - Canadian Centre for Cyber Security](#)
- Artificial Intelligence: [Artificial Intelligence - ITSAP.00.040 - Canadian Centre for Cyber Security](#)
- Potential Threat Implications of Generative AI for Corporate Security published 18th July 2023 by FBI
- [Generative AI and Data Privacy: A Primer, Congressional Research Service, 23rd May, 2023](#)
- <https://learn.microsoft.com/en-us/microsoft-365-copilot/microsoft-365-copilot-privacy#how-does-microsoft-copilot-for-microsoft-365-protect-organizational-data>
- Microsoft is fully committed to the Responsible AI standards: [Responsible AI Principles and Approach | Microsoft AI](#)
- On information accuracy decision record stating that users are responsible to verify accuracy: "Can I trust that Copilot's answers are always accurate?" [Frequently asked questions about Microsoft 365 Copilot](#)
- [Semantic Index for Copilot: Explained by Microsoft](#)
- [Connect to the Microsoft Copilot Dashboard \(Preview\) | Microsoft Learn](#)
- [Overview of Copilot | Microsoft Learn](#)
- [Copilot Privacy and Protections | Microsoft Learn](#)