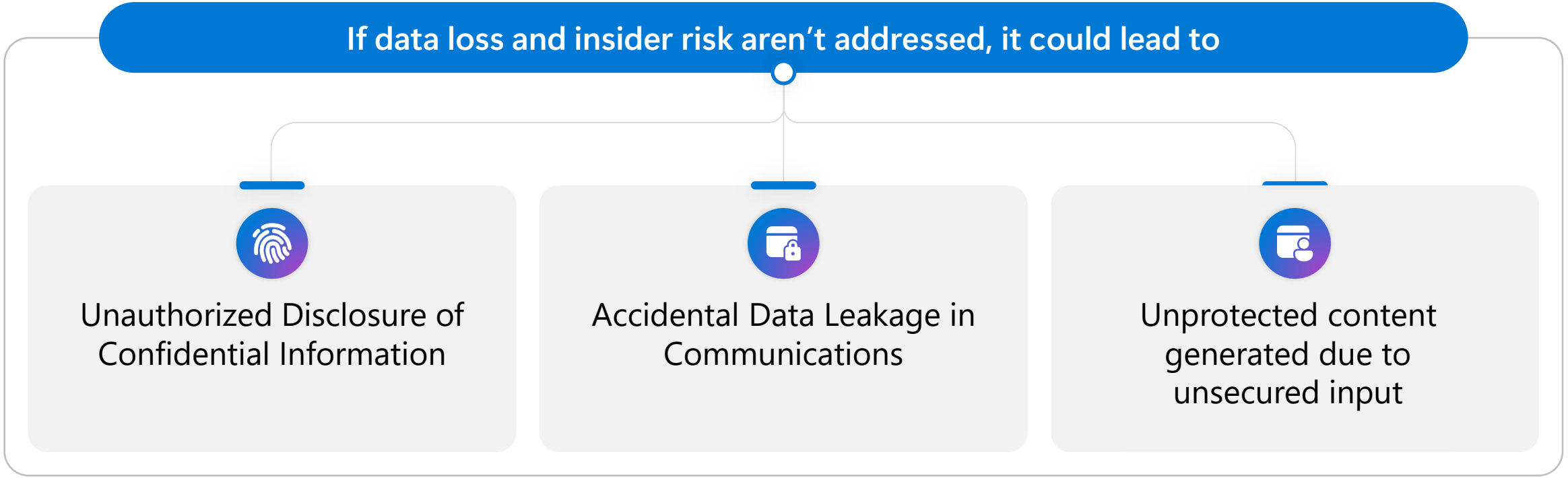
Microsoft

Protect Against Data Loss & Insider Risk for M365 Copilot

Microsoft Deployment Blueprint

Problem Summary

Microsoft 365 Copilot (Copilot)'s ability to leverage information available to employees has raised concerns for organizations about data loss and insider risk.



By addressing these risks, organizations can ensure that information is shared appropriately and securely

Common causes of Data loss & Insider risk

What is data loss and insider risk?

Data loss and insider risks may occur without Copilot use. The productivity benefits of these tools create opportunities for these risks to happen with less effort

Data loss is when unsafe or inappropriate sharing, transfer, or use of sensitive content occurs

Insider risk involves users misusing their authorized access to cause harm. This misuse may be intentional or unintentional

What causes data loss and insider risks?

Copilot creates documents that aren't protected by default, because the inputs weren't protected

Using Copilot to access or summarize sensitive data in a way that violates usage guidelines

A user's credentials are compromised, and the bad actor uses Copilot to access sensitive data

An employee has resigned, wants to keep company IP, and uses Copilot

Protect Against Data Loss and Insider Risk for M365 Copilot

	Establish Visibility and Risk Awareness	Apply Protections to Copilot Interactions	Monitor and Respond to Risky Behavior
Activities	• Enable and Review DSPM for AI Insights • Conduct a preliminary review in Activity Explorer • Align User Risk Levels with Organization Risk Appetite	• Establish Sensitivity Labeling and Permissions Framework • Assess Sensitive Data Landscape • Turn on DLP for Copilot to protect sensitive content	• Audit risky behavior with Communication Compliance • Establish Risk-Based Controls with Insider Risk Management • Restrict high risk users from interacting with sensitive data
Outcomes	Establish visibility into usage and data exposure to inform risk-based protection strategies.	Ensure Copilot-generated content is governed and protected by default	Detect and mitigate insider threats and misuse of Copilot
Effort*	2–4 days	1-2 weeks	1-2 weeks

*Suggested efforts should be reviewed into timelines based on your tenant size and organizational complexity

Establish Visibility and Risk Awareness

1. Enable and Review DSPM for AI Insights
2. Conduct a preliminary review in Activity Explorer
3. Align User Risk Levels with Organization Risk Appetite
4. Monitor Agent Growth and Security Posture

Outcome

Establish visibility into usage and data exposure to inform risk-based protection strategies.

Enable and Review DSPM for AI Insights

1. Enable DSPM for AI and foundational policies
   - Ensure Microsoft Purview Audit is turned on to capture user activity
   - Enable the IRM Data Leaks and Communication Compliance default Microsoft Purview DSPM for AI policies to begin monitoring Copilot interactions and data exposure.

2. Review DSPM for AI Reports
   - **Copilot Usage Overview:** Use the Copilot usage chart to assess how frequently users are interacting with Copilot. High usage may indicate a need to accelerate security policy implementation to mitigate potential risks.
   - **Sensitive Information Types (SITs) in Copilot:** Examine which SITs are involved in Copilot interactions. Identify any unexpected SITs and determine whether additional monitoring or restrictions are needed.
   - **Unethical Interactions:** Review flagged unethical interactions based on Communication Compliance (CC) policies and associated trainable classifiers. This helps assess current misuse and informs future policy creation.

Establish Visibility and Risk Awareness

1. Enable and Review DSPM for AI Insights
2. Conduct a preliminary review in Activity Explorer
3. Align User Risk Levels with Organization Risk Appetite
4. Monitor Agent Growth and Security Posture

Outcome

Establish visibility into usage and data exposure to inform risk-based protection strategies.

Enable and Review DSPM for AI Insights

2. Review DSPM for AI Reports
    - **User Risk Levels:** Evaluate the risk level of users interacting with Copilot, as determined by the IRM policy. High-risk users should be considered for inclusion in adaptive protection policies and monitored closely (see Phase 3 guidance).
    - **Sensitivity Labels in Use:** Identify the top sensitivity labels referenced in Copilot interactions. Pay special attention to:
        - Unlabeled files being accessed by Copilot, which may indicate unprotected data exposure.
        - Confidential or highly sensitive labeled files being accessed, which may exceed your organization's risk tolerance.

Establish Visibility and Risk Awareness

1. Enable and Review DSPM for AI Insights
2. Conduct a preliminary review in Activity Explorer
3. Align User Risk Levels with Organization Risk Appetite
4. Monitor Agent Growth and Security Posture

Outcome

Establish visibility into usage and data exposure to inform risk-based protection strategies.

Conduct a Preliminary Review in Activity Explorer

1. Begin with a preliminary review of Copilot interactions using DSPM for AI Activity Explorer
   - Filter by Copilot or specific agents to view interaction-level activity details.
   - Review the actual prompt and response text to understand the nature of user queries and Copilot's replies.
   - Apply filters to prioritize reviewing activities involving:
     - High-risk users (as defined by IRM policies)
     - Most sensitive information types (SITs)
     - Highly confidential sensitivity labels
     - You will also be able to view **web query data**, which shows what information is being accessed or transmitted beyond your tenant boundary.

Establish Visibility and Risk Awareness

1. Enable and Review DSPM for AI Insights
2. Conduct a preliminary review in Activity Explorer
3. Align User Risk Levels with Organization Risk Appetite
4. Monitor Agent Growth and Security Posture

Outcome

Establish visibility into usage and data exposure to inform risk-based protection strategies.

Align User Risk Levels with Organization Risk Appetite

1. Adjust user risk levels in IRM to reflect your organization's definition of risky behavior
   - Navigate to Microsoft Purview Insider Risk Management (IRM).
   - Review and calibrate user risk levels based on:
     - Your organization's risk tolerance.
     - Behavioral indicators that your organization considers risky (e.g., frequent access to sensitive data, unusual Copilot usage patterns).
   - This ensures that high-risk users are accurately identified and prioritized for monitoring and protection policies in later phases.
   - For more information on Insider Risk settings to understand and choose the settings that best meet the Compliance needs for your organization, please see: Learn about Insider Risk Management Settings

Establish Visibility and Risk Awareness

1. Enable and Review DSPM for AI Insights
2. Conduct a preliminary review in Activity Explorer
3. Align User Risk Levels with Organization Risk Appetite
4. Monitor Agent Growth and Security Posture

Outcome

Establish visibility into usage and data exposure to inform risk-based protection strategies.

Monitor Agent Growth and Security Posture

1. Review Copilot and Copilot Studio agent trends in the Apps and Agents page
   - Track usage trends for Copilot and Copilot Studio agents to understand adoption velocity. Rapid growth may signal the need to assess associated data risks.
   - Investigate whether these agents are interacting with sensitive data by correlating usage with sensitivity labels and SITs.
   - Evaluate the security posture of each app or agent:
     - Confirm whether appropriate protection policies are already applied.
     - Identify any gaps in policy coverage that may require remediation.
2. If there are AI regulations you need to adhere to, review Compliance Manager
   - Build an assessment in Purview Compliance Manager for one of the following AI-related regulations:
     - EU Artificial Intelligence Act
     - ISO/IEC 23894:2023
     - ISO/IEC 42001:2023
     - NIST AI Risk Management Framework (RMF) 1.0
   - These assessment will help to take inventory of any data protection risks you may have when staying current with regulations and certifications.
   - Please see our documentation for more details: Assessments for AI Regulations

Apply Protections to Copilot Interactions

1. Establish Sensitivity Labeling and Permissions Framework
2. Assess Sensitive Data Landscape
3. Turn on DLP for Copilot to protect sensitive content

Outcome

Ensure Copilot-generated content is governed and protected by default

Establish Sensitivity Labeling and Permissions Framework

1. Go to Microsoft Purview Information Protection
   - Ensure Microsoft Information Protection (MIP) labeling is in place by publishing sensitivity labels to Office apps, SharePoint sites, and Outlook.
     - This enables users to manually apply data protection and builds familiarity with labeling during Copilot or Copilot agents deployment.
     - **Labeling is foundational to Copilot's security model:** Copilot adheres to user permissions defined by sensitivity labels, ensuring it only accesses content users are authorized to view. Labels are inherited in Copilot responses, and visual indicators help users stay aware of data sensitivity.
     - Set up auto-labeling policies to automatically apply sensitivity labels to files containing Sensitive Information Types (SITs). You can configure these labels to enforce access controls—granting or restricting access based on user roles or group membership.
     - Refer to the Secure by Default Blueprint for more details on sensitivity labeling in a Copilot deployment.
   - If labeling is already in place, evaluate the current state of labels to determine if adjustments need to be made. Ensure there is no more than a 5x5 label scheme. Will also need to evaluate the settings for each label to review the location scope, encryption settings, and the user/group permissions.

Apply Protections to Copilot Interactions

1. Establish Sensitivity Labeling and Permissions Framework
2. Assess Sensitive Data Landscape
3. Turn on DLP for Copilot to protect sensitive content

Outcome

Ensure Copilot-generated content is governed and protected by default

Assess Sensitive Data Landscape

1. Go to Microsoft Purview Information Protection > Explorers > Data Explorer
   - First, determine the highest priority sensitive information.
   - Under Sensitive Information Types, select one of those sensitive information types. This will show you where the sensitive information is in the organization as well as which of this content has been labeled with a sensitivity label.
   - Understanding where sensitive information is within your environment and if it is protected by a sensitivity label will help you determine who should be able to have access to this content using Copilot and Agents and what protections this content should have.

Apply Protections to Copilot Interactions

1. Establish Sensitivity Labeling and Permissions Framework
2. Assess Sensitive Data Landscape
3. Turn on DLP for Copilot to protect sensitive content

Outcome

Ensure Copilot-generated content is governed and protected by default

Turn on DLP for Copilot to protect sensitive content

1. In Microsoft Purview, go to the Data Loss Prevention (DLP) section and create a new DLP for Copilot policy.
    - Configure a DLP policy to target files that have been labeled using Microsoft Purview Information Protection.
    - Enable the setting to block Copilot and Copilot Studio agents from accessing these labeled files. This will:
        - Prevent Copilot from generating responses based on sensitivity labeled files containing specific SITs.
        - Block summarization of such files if it violates usage guidelines.
        - Restrict Copilot from processing emails labeled with sensitivity tags (e.g., Confidential, Highly Confidential). This will ensure Agents with the capabilities to send emails will not have access to sensitivity labeled mail.

Monitor and Detect Risky Behavior

1. Audit risky behavior with Communication Compliance
2. Establish Risk-Based Controls with Insider Risk Management
3. Restrict high risk users from interacting with sensitive data

Outcome

Detect and mitigate insider threats and misuse of Copilot

Audit risky behavior with Communication Compliance

1. Navigate to Communication Compliance (CC)
   - Create a policy to detect prompt injection attacks or other risky behaviors while using Copilot. Use this to audit user activity and identify potential misuse.
2. If you need to monitor and respond for legal reasons, please note that eDiscovery is available to assist with these requirements.
   - Preserve, collect, review, analyze, and export data relevant to investigations.
   - Manage custodians, legal holds, review sets, and case content.
   - Use predictive coding and attorney-client privilege detection to streamline legal review.
   - Please refer to the provided documentation for further guidance.

Monitor and Detect Risky Behavior

1. Audit risky behavior with Communication Compliance
2. Establish Risk-Based Controls with Insider Risk Management
3. Restrict high risk users from interacting with sensitive data

Outcome

Detect and mitigate insider threats and misuse of Copilot

Establish Risk-Based Controls with Insider Risk Management

1. Ensure Insider Risk Management (IRM) policies are in place
    - Enable Adaptive Protection within the Adaptive Protection settings tab. Adaptive Protection enables security administrators to dynamically safeguard their organizations by automatically applying stricter security policies to users based on their risk level, eliminating the need for manual intervention.

Monitor and Detect Risky Behavior

1. Audit risky behavior with Communication Compliance
2. Establish Risk-Based Controls with Insider Risk Management
3. Restrict high risk users from interacting with sensitive data

Outcome

Detect and mitigate insider threats and misuse of Copilot

Restrict high risk users from interacting with sensitive data

1. Set up Adaptive Protection policies to restrict high-risk users from risky actions

2. Restrict access to sensitive sites:
    - In Microsoft Entra ID (formerly Azure AD), create a Conditional Access policy:
        - **Conditions:**
            - User risk level: Elevated
            - Target resources: SharePoint Online, OneDrive, Teams, or specific sensitive sites
        - **Access controls:**
            - Block access or require MFA
            - Optionally apply session controls to limit access

3. Prevent data deletion:
    - Use IRM data leak policies to assign elevated risk when triggered.
    - Configure a Data Lifecycle Management (DLM) policy to preserve deleted files for elevated-risk users.
    - Create a DLP policy with the condition: “User’s risk level for Adaptive Protection is Elevated.” Set the action to block deletion or apply audit-only mode for lower-risk users.

Monitor and Detect Risky Behavior

1. Audit risky behavior with Communication Compliance
2. Establish Risk-Based Controls with Insider Risk Management
3. Restrict high risk users from interacting with sensitive data

Outcome

Detect and mitigate insider threats and misuse of Copilot

Restrict high risk users from interacting with sensitive data (continued)

4. Prevent data downloads:
    - In the Microsoft Purview Compliance Portal, create a DLP policy:
        - **Condition:** User's risk level is Elevated
        - **Action:** Block downloads from SharePoint, OneDrive, Teams, Exchange, or endpoint devices
        - **Optional:** Show a policy tip or alert
        - **Scope:** Apply to Exchange Online, Teams, Devices

5. Prevent printing of sensitive data:
    - Create a DLP policy to block printing:
        - **Condition:** User's risk level is Elevated
        - **Action:** Block printing from SharePoint, OneDrive, Teams, Exchange, or endpoint devices
        - **Scope:** Devices and SharePoint Online

6. Monitor effectiveness using Activity Explorer and Insider Risk dashboards. Adjust thresholds or actions as needed.

If Data Loss and Insider Risk isn't addressed, they could lead to:

Disclosure of Confidential Information and Data Leakage.

By addressing these risks through the following actions, organizations can ensure that information is shared appropriately and securely.

Phase 1: Establish Visibility and Risk Awareness

Understand usage and exposure to inform protection strategies.

- Enable & review DSPM for AI reports
- Conduct preliminary review in Activity Explorer
- Align user risk levels with org risk appetite
- Monitor agent growth and security posture
- Build AI regulation assessments in Compliance Manager

Phase 2: Apply Protections to Copilot Interactions

Apply protections to ensure Copilot respects data boundaries.

- Establish sensitivity labeling and permissions framework
- Turn on DLP for Copilot to block access to sensitive content
- Apply Conditional Access for elevated-risk users
- Prevent risky actions: printing, downloading, deleting
- Automate labeling and enforce access controls

Phase 3: Monitor and Respond to Risky Behavior

Detect and respond to misuse or insider threats.

- Audit behavior with Communication Compliance
- Enable Adaptive Protection for high-risk users
- Use eDiscovery for legal investigations
- Restrict high-risk users from sensitive actions/sites
- Monitor effectiveness and refine thresholds

Microsoft

Thank you

Microsoft Deployment models

Read the detailed guide for this model at aka.ms/Copilot/OversharingBlueprintLearn