Mitigate Copilot for Microsoft 365 Risks Through Information Governance

Published 17 November 2023 - ID G00800669 - 60 min read

By Analyst(s): Max Goss, Larry Cannell

Initiatives: Digital Workplace and CRM for Technical Professionals; Security Technology and Infrastructure for Technical Professionals

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.

Overview

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.
- 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.

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 Purview Information Protection, and Data Loss Prevention. This reduces the risk of oversharing and accidental discovery. Before deploying Copilot, test your access controls using search.
- 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.
- 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.

Analysis

Microsoft Copilot for Microsoft 365 (Copilot) promises a radical shift in terms of how users find, create, consume and understand content. However, its impact and value are highly dependent on the organizational information stored in Microsoft 365. For many organizations, information governance in Microsoft 365 is immature, and many struggle to manage the risks of overshared and poorly permissioned content. According to Varonis’s “The Great SaaS Data Exposure” report (2022), for the average organization using Microsoft 365, 1 in 10 sensitive files is exposed to every user. [1]

Copilot makes information much easier to find. If no action is taken to improve information governance before it is deployed, it could significantly increase the risks that poorly permissioned and/or redundant or out-of-date information will be exposed.

This research outlines the steps you need to take to improve your information governance to get the most out of Copilot and to mitigate the associated risks.

It is split into four main sections:

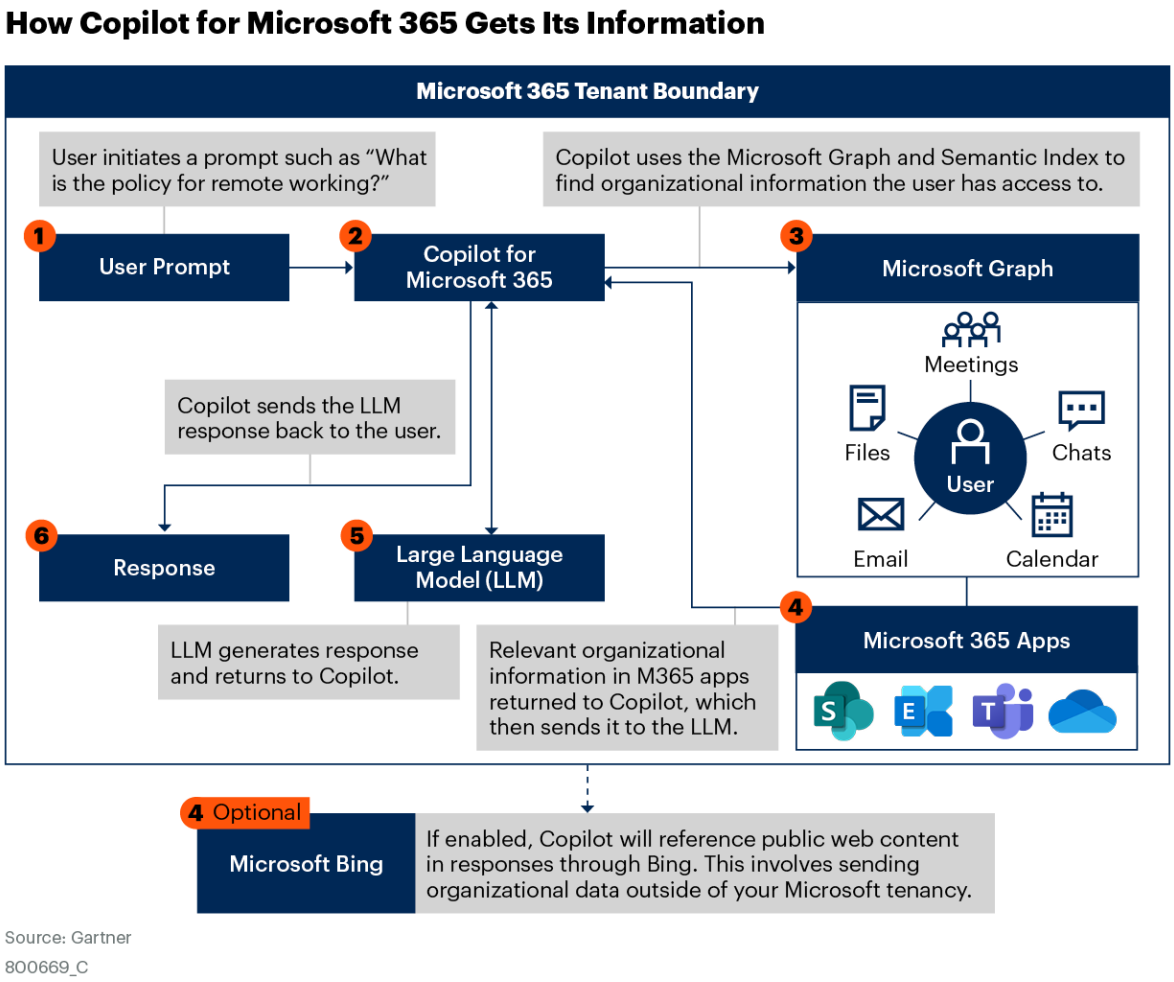
How Copilot for Microsoft 365 Works With Your Organization's Information

Back

When compared with other generative AI tools, the unique value proposition of Copilot for Microsoft 365 is that it can work directly with the data stored in your organizations' Microsoft 365 tenant. Copilot can summarize meetings in Teams, create draft content in Word and PowerPoint, and analyze data in Excel. It can also be used to create new SharePoint sites, Power Apps and Power Automate flows, all through a simple natural language prompt. Additionally, Copilot has a Microsoft 365 Chat ("M365 Chat") function, which is integrated into Microsoft Teams and available as a stand-alone Microsoft 365 application. This can be used to discover, summarize and reason over data across your Microsoft 365 tenant. Copilot differs significantly to other well-known generative AI services (such as OpenAI's ChatGPT, Bing Chat and Bing Chat Enterprise) as these tools work on publicly available data (unless organizational content is added through a prompt or to their underlying model).

Copilot for Microsoft 365 accesses organizational information through Microsoft Graph, which is the API and data layer behind Microsoft 365. It can return and synthesize information from documents, pages, emails, chats and meetings where the user has "view" access or above. It uses the Microsoft 365 Semantic Index to determine context and improve relevance before sending the organizational data to the large language model (LLM) to formulate a response back to the user. Figure 1 shows a simplified version of how this works.

Figure 1: How Copilot for Microsoft 365 Gets Its Information



In this example, a user asks Copilot for information using Copilot's chat interface in Microsoft Teams.

1. A user starts a chat with Copilot and asks, "What is the policy for remote working?"

2. The prompt (the user query) is sent to the Copilot orchestration engine for processing.

3. The orchestration engine queries Microsoft Graph looking for relevant information (that the user has access to) which could match their query.

4. Copilot uses the Semantic Index, which helps to provide context and ensure relevance when retrieving information from the underlying Microsoft 365 apps (including SharePoint, OneDrive and Exchange) using search. The search results are security trimmed so that only information that the user has access to is returned.

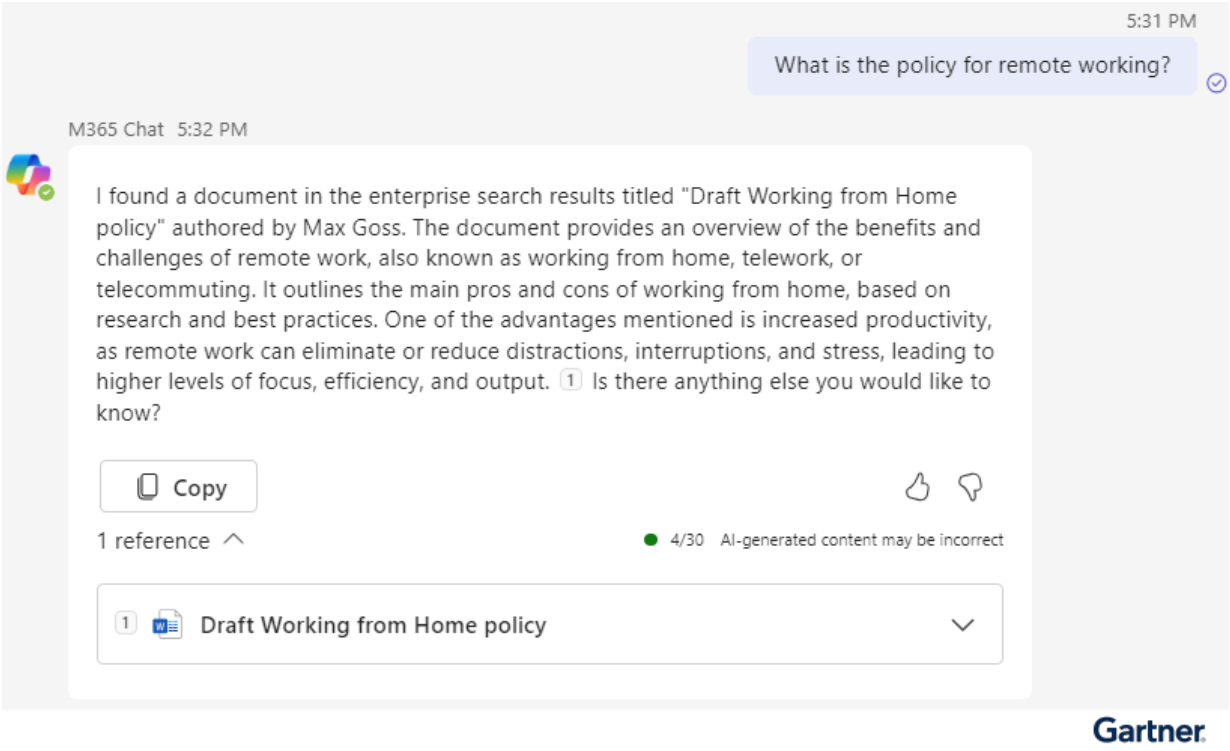
5. The user prompt is augmented (grounded) with the discovered information (documents or messages, for example) and then sent to the LLM, which generates a response.

6. The Copilot orchestration engine delivers the response back to the end user.

Note: If enabled at the tenant level, Copilot for Microsoft 365 can choose to search the internet through Bing to use up to date web content to improve the response. This is depicted in the graphic in optional step 4 and can involve sending organizational data outside of your tenancy. The implications of this are discussed in the next section.

In our example, the response will be delivered via the Copilot chat interface in Microsoft Teams (M365 Chat). The response might include a summary of relevant work-from-home policy documents that were identified through search. Copilot will also list its sources. Figure 2 shows an example response.

Figure 2: Copilot Response



Is Your Organization's Information Safe With Copilot for Microsoft 365?

Microsoft asserts that "*Prompts, responses and data accessed through Microsoft Graph aren't used to train foundation LLM's including those used by Copilot for Microsoft 365.*" Microsoft also states that all data remains within the Microsoft 365 service boundary and that data residency rules are respected. When processing data, they note that *"Copilot for Microsoft 365 calls to the LLM are routed to the closest data centers in the region, but also can call into other regions where capacity is available during high utilization periods."* Microsoft asserts that all EU traffic stays within the EU data boundary but *"worldwide traffic can be sent to the EU and other countries or regions for LLM processing*." [2]

Copilot for Microsoft 365 uses Microsoft Azure OpenAI services hosted within the Microsoft Azure environment for LLM processing. It does not interact with or share data with publicly available services operated by Open AI (such as ChatGPT). Microsoft states that customer prompts and responses are not available to other Microsoft 365 organizations or to OpenAI. [3]

If your organization has strict data residency, privacy or sovereignty requirements, Gartner strongly recommends that you work with Microsoft and your compliance team to fully understand the end-to-end information flow before rolling out Copilot. Microsoft provides more information on data privacy here: Data, Privacy, and Security for Copilot for Microsoft 365, Microsoft Learn.

Copilot for Microsoft 365 respects user permissions. It only returns information stored in Microsoft 365 services (such as OneDrive, Teams, SharePoint and Outlook mailboxes) that the user has access to. Additionally, Copilot will only search across the user's home tenant, and will not search content across tenants the user may have access to as a B2B user. Content is only discoverable to Copilot if it has been indexed by Microsoft Search. If a user can find something in search, it is highly likely that they will find it through Copilot.

How Copilot Works With Public Web Content

Copilot's underlying LLM was trained on public web data up until September 2021. Like ChatGPT, the model does not know about events that have happened since that date. Ask Copilot, "Who won Wimbledon in 2023?" and the LLM cannot provide an answer. To get around this limitation, the M365 Chat experiences in Copilot have an integration with Bing, which is enabled by default. If Copilot determines that current web content will improve the quality of the response, it will generate a search query to Bing. Using the results from the Bing search engine, it can then provide up-to-date answers. Users must enable the "external links" toggle In Copilot's M365 Chat interface before Copilot can use Bing's search. They can only do this if the setting is enabled at the tenant level (which it *is* by default).

When Copilot sends a search query to Bing, the user and tenant ID are not included in the search. However, organizational data that the user can access will likely be included. This data is sent outside of your Microsoft 365 tenancy.

The Microsoft Products and Services Data Protection Addendum (DPA) does not apply to the use of Bing.

Microsoft asserts that "*Microsoft Bing is a separate service from Microsoft 365 and data is managed independently from Microsoft 365.*" Bing is not covered by Microsoft's EU data boundary commitment nor any agreement that requires company data to remain in a geographic area or tenant boundary. This is significant. If you accept the default settings, and allow M365 Chat experiences in Copilot to search Bing for current web content, there is a risk that sensitive organizational data could leave your tenancy. It also means that you lose control of where your data is being processed. For regulated organizations with strict data residency requirements, Gartner urges caution. You must consult with your legal and compliance teams to determine if this risk is acceptable. Copilot can still reference public data if this integration is disabled, but it will be limited to when the underlying LLM was last updated (currently September 2021). Note, this setting does not affect Copilot's ability to interact with your internal organizational data (see Manage Access to Public Web Content in Copilot for Microsoft 365 Responses).

Avoiding Another "Delve Moment"

Despite the safeguards, Copilot for Microsoft 365 can still expose organizations to risk.

When Microsoft released Office Graph (which evolved into Microsoft Graph) and Delve in 2014, Gartner clients raised concerns when Delve recommended content to users that they should not have access to. Delve uses Microsoft Graph to proactively recommend content to users based on a multitude of factors, such as what they have recently been working on and who they have been working with. Like Copilot, Delve respects user permissions and only promotes content that the user has access to.

When users reported that Delve was displaying sensitive information that they should not have access to, IT faced strong calls to "switch off" Delve. For some organizations, the Delve brand became synonymous with oversharing and data leakage. While Delve was technically not the source of the problem, it did shine a light on content that had been overshared or stored in the wrong location, — for example sensitive content stored in a public group or team. In other words, the root cause of the "Delve moment" was poor information governance.

Copilot for Microsoft 365 could create "another Delve moment." However, this time it will be worse, as Copilot will construct and formulate answers based on content that the user has access to, rather than just display a list of documents.

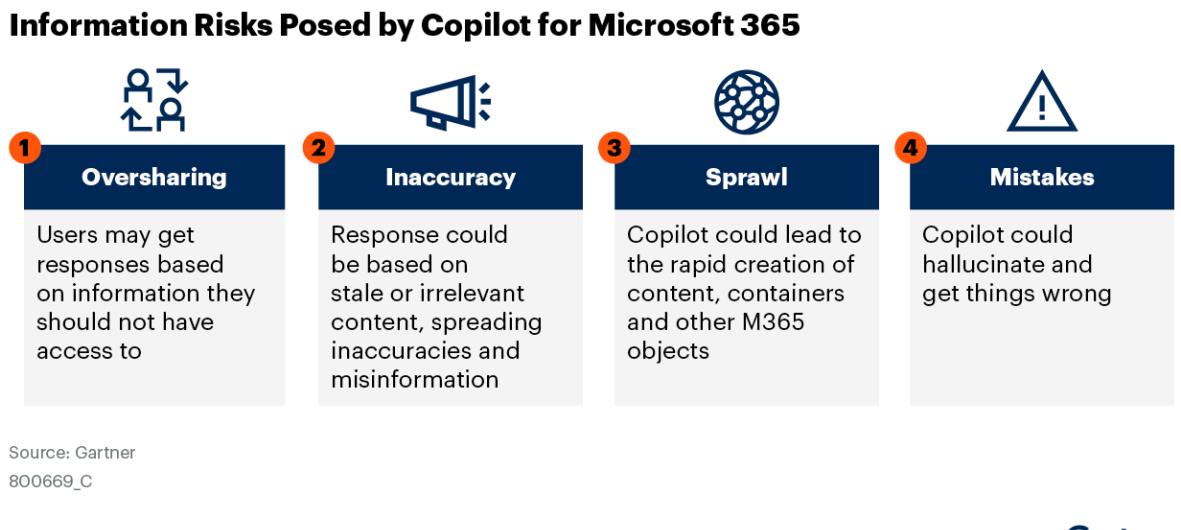
If the underlying permissions are incorrect, or if content is stale or redundant, there is a strong probability that users will either receive information that they should not have access to, or be given wrong or misleading responses. The following sections explore these information risks in detail, and provide guidance on how to mitigate them.

Information Risks Posed by Copilot for Microsoft 365

Back

Gartner has identified four key risks that Copilot for Microsoft 365 could either cause or accentuate (see Figure 3).

Figure 3: Information Risks Posed by Copilot for Microsoft 365



To further understand some of these risks, Gartner carried out a series of tests using a preview version of Copilot for Microsoft 365 on a demo tenant. Most of these tests were performed through the M365 Chat interface in Microsoft Teams. It should be highlighted that the tenant contained a limited data set with a preview version of Copilot, and so any findings should be seen as illustrative only. While Copilot fully respects user permissions and Microsoft 365 access controls, our testing did highlight that poor information governance can result in Copilot oversharing or returning incorrect information. As such, organizations must prepare for Copilot by improving their information governance to mitigate risk and ensure that they get best value out of the product.

Oversharing

In Gartner's 2023 Microsoft 365 survey, [4] 59% of respondents cited oversharing and data loss as among the three biggest risks to their organization's Microsoft 365 deployment. Oversharing issues in Microsoft 365 and how to mitigate them is also a common topic during Gartner inquiry calls. Many users do not know how to correctly store and share content in Microsoft 365. Gartner often sees examples of users sharing sensitive files using the "Everyone in the Company link," or putting confidential information in public groups or teams.

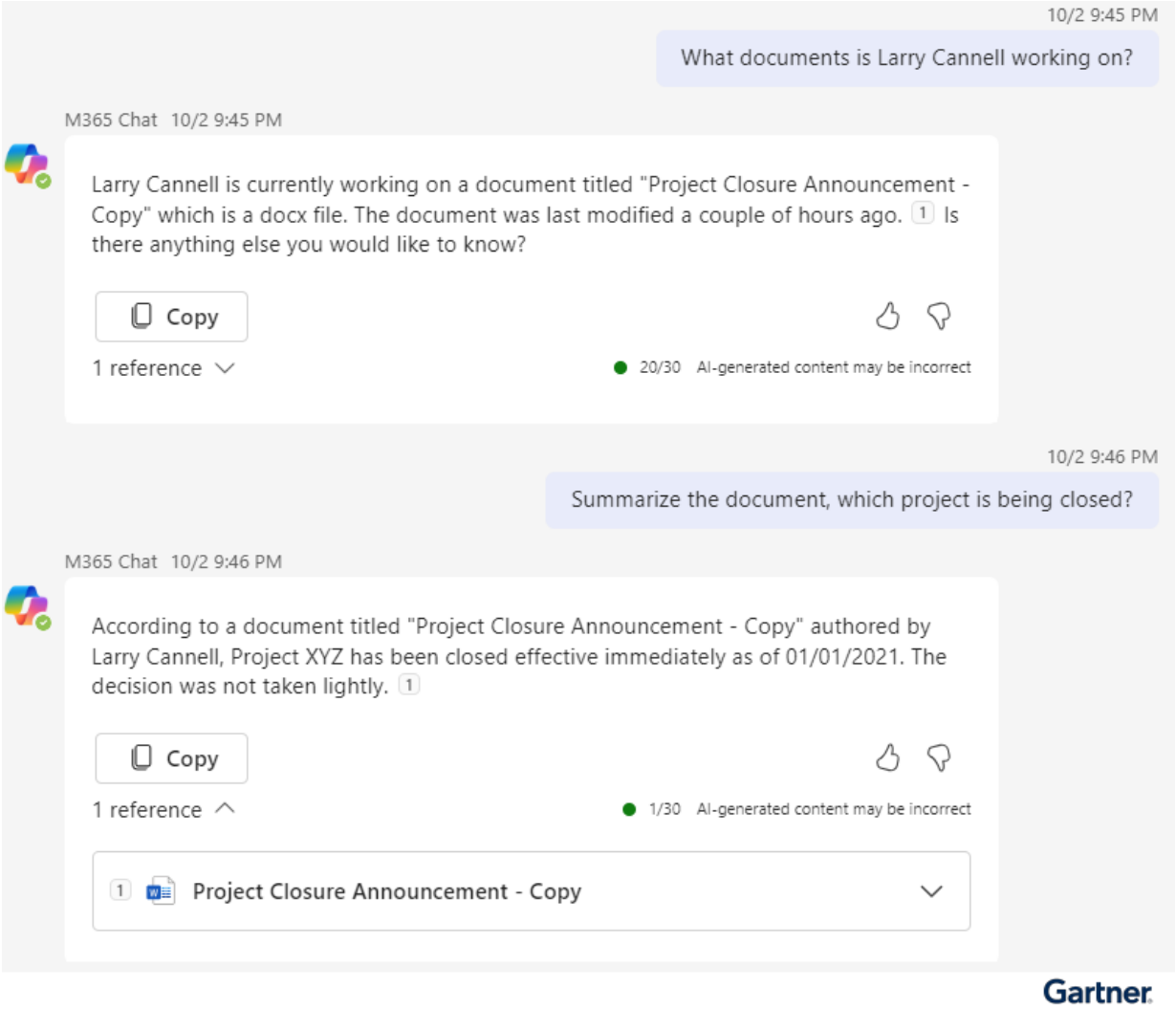
In a November Gartner 2023 webinar poll, when asked "What do you think is the biggest risk/issue with deploying Copilot for Microsoft 365," the top risk was "Oversharing and exposing sensitive information to employees." [5]

Working in the test environment, we created a scenario where an executive called "Larry Cannell" was working to close down a stalled project. In this test, he put a Word document about the planned project closure into a team in Microsoft Teams. The team had been incorrectly set to "public," meaning that its contents were accessible to all members of the organization, if they knew where to look.

Logging in as another user who had never accessed the public team nor had worked on the document, we asked Copilot through the M365 Chat interface in Microsoft Teams "*What documents is Larry Cannell working on*?" Copilot responded with a link to the project closure document. We then asked it to summarize the document and tell us which project was being closed. Copilot proceeded to provide a brief summary naming the project and saying, "*The decision was not taken lightly*"!

Figure 4 shows the exchange in full.

Figure 4: Copilot for Microsoft 365 Oversharing Example



The more powerful the information retrieval tool, the more critical it is to have effective information governance controls and permissions management.

Copilot for Microsoft 365 can be used as a powerful information retrieval tool and should result in significant productivity benefits in terms of making information easier to find. However, with this innovation, there is also risk that needs to be managed. In our mocked-up example, a user was able to access potentially very sensitive information just by knowing the name of an executive. While they may have found the same document through search, Copilot's natural language interface has made it easier. Additionally, the summarization may reveal important information that, when taken out of context (the user may not read the whole document), could be very damaging. With Copilot, users who overshare sensitive information or store it in the wrong place are more likely to expose it to others.

Inaccuracy

Many organizations struggle with their content life cycle, and do not have effective retention policies to handle stale or redundant content. This has always been a problem for search. Even if permissions are set correctly, finding and using out-of-date content such as old policies or pricing information can lead to embarrassment, compliance issues and reputational damage. Indeed, regulated organizations may struggle to satisfy their compliance requirements if they cannot prove that content was deleted after a certain period.

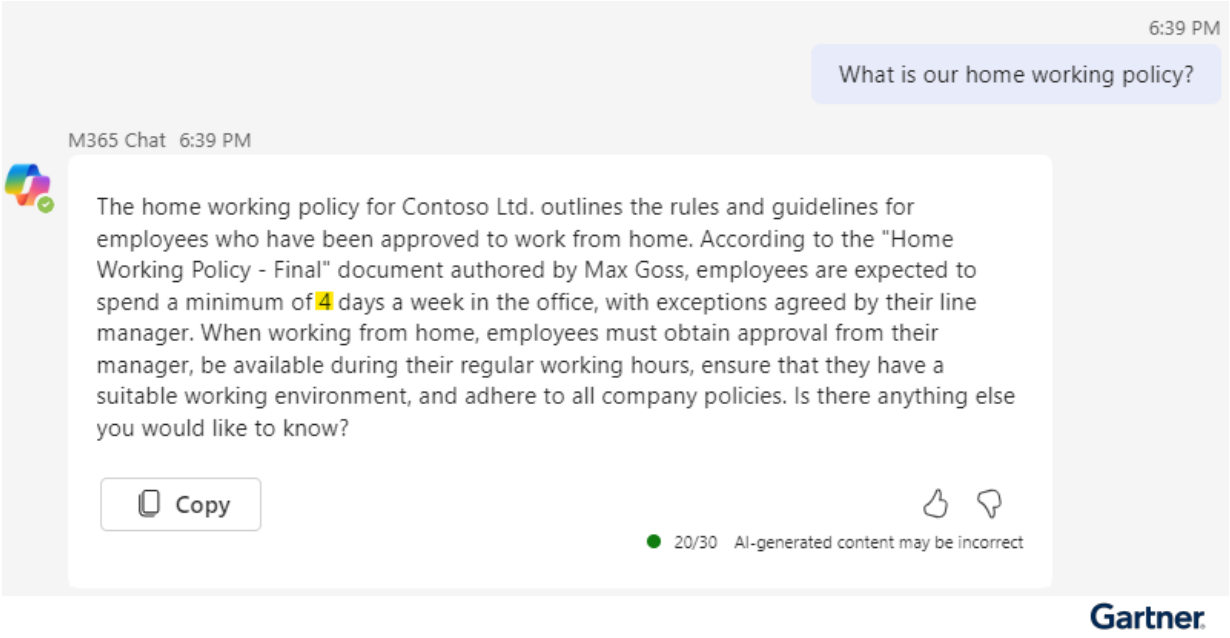
> **As most organizations do not have a well-managed content estate, the chances of Copilot for Microsoft 365 finding or synthesizing redundant, obsolete or trivial (ROT) information is high.**

Copilot benefits from the same relevance features as Microsoft Search, which includes recognizing items that are frequently clicked on, viewed in the tenant and previously viewed by the user. However, it cannot be specifically taught to focus on authoritative content, for example, choosing content in one site over another.

In the test environment, we created a scenario where two home-working policies were created. One was uploaded to the organization's HR-Policies SharePoint site, which is the authoritative source for all HR policies. The other was uploaded to a user's OneDrive and then shared with a wide group of users. The policies were almost identical except the one on the HR-Policies site stated correctly that employees are expected to spend a "minimum of 3 days a week in the office." The policy in OneDrive stated incorrectly that employees are expected to spend a "minimum of 4 days a week in the office" The OneDrive version was designed to simulate a work-in-progress draft that had been shared for review.

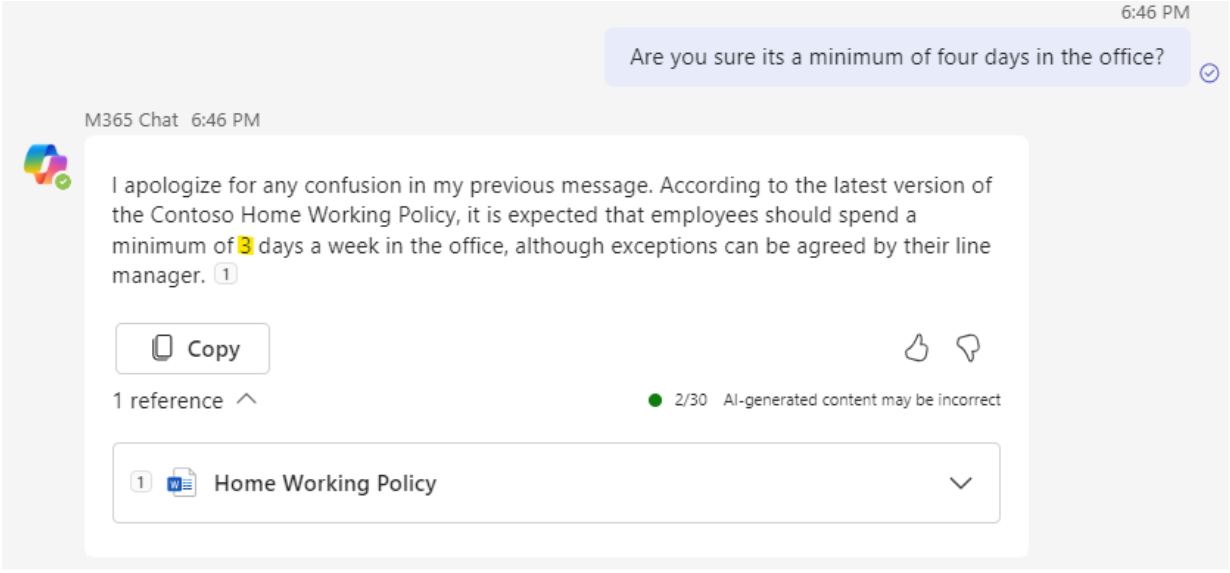
Using another user account that had access to both files, we asked Copilot through the M365 Chat interface in Microsoft Teams, "*What is our home working policy?*" Copilot found the policy stored in the original user's OneDrive and summarized it stating that employees are expected to spend a "minimum of 4 days a week in the office." It did not return the version of the policy on the HR SharePoint site. Figure 5 shows the prompt and response in full. Interestingly, in this test, Copilot did not provide a link to its source.

Figure 5: Copilot for Microsoft 365 Inaccuracy Example



We continued to probe Copilot, by asking it "*Are you sure it's a minimum of 4 days in the office?*" At this point, Copilot changed its answer and returned the correct policy, apologizing for the confusion and noting that "*employees should spend a minimum of 3 days a week in the office.*" However, a different user asking the same question continued to get the "4 day" response. Figure 6 shows Copilot's corrected response.

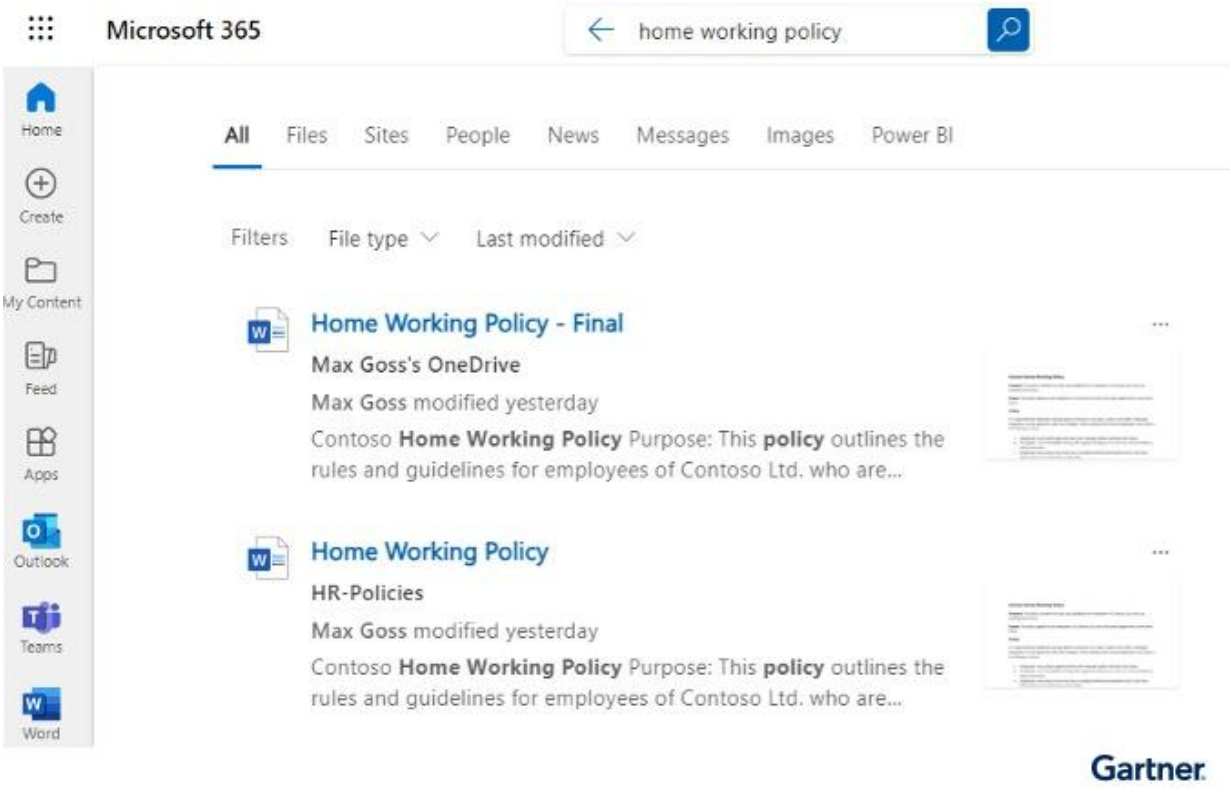
Figure 6: Copilot for Microsoft 365 Correcting Itself



Poor information governance — in this case having two copies of the same policy that can be accessed by end users — will cause confusion and potentially spread misinformation. In Microsoft search, you can partially mitigate this risk by specifying promoted results (or bookmarks) which will always return based on particular key words or phrases. Copilot for Microsoft 365 does not yet have this functionality.

If you retain redundant and obsolete data, there is a risk that Copilot or search will find it. However, because Copilot will summarize the content into a confidently written answer, the original source may not be immediately obvious. For example, in search results, we can see that there are two home working policies: one stored on a user's OneDrive, the other on the HR-Policies SharePoint site (see Figure 7). In Copilot, this is obscured.

Figure 7: Search Results for "Home Working Policy"



We ran a second test where we asked Copilot "*What is our policy on records management?*" Copilot successfully found and summarized the records management policy that we had uploaded to the HR-Policies site. However, it left out a key part of the document that said it was only applicable to HR professionals, and that other departments should consult locally. The potential for generative AI to confidently spread misinformation within an organization is huge. As well as taking sensible steps to improve permissions, retention and your information architecture, users must be educated to adopt a "distrust and verify" approach to AI-generated content and summarizations.

Sprawl

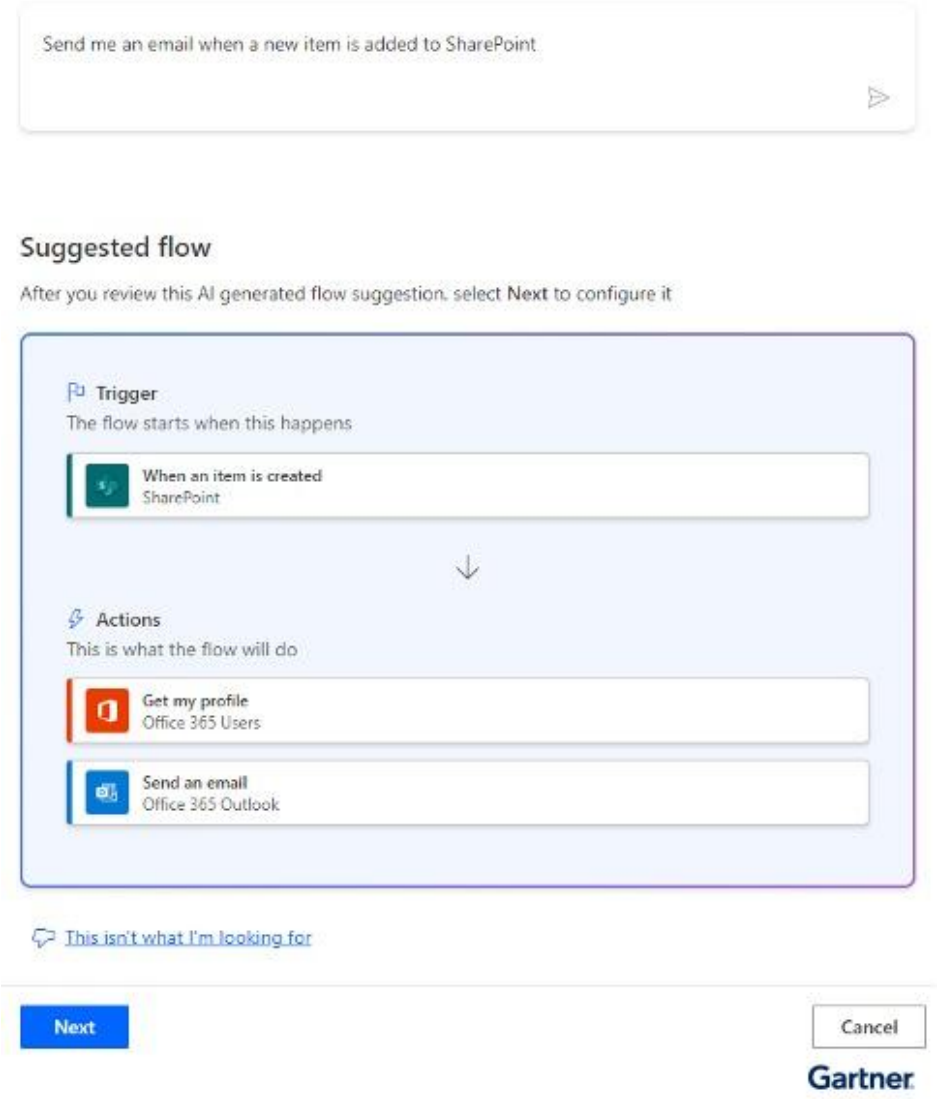
Microsoft asserts that, on an average workday, their customers add over 2 billion new files to Microsoft 365. Digital content is growing rapidly and concerns about content sprawl are top of mind for many organizations. In Gartner's 2023 Microsoft 365 survey, 58% of respondents cited content sprawl as among the biggest risks to their organization's Microsoft 365 deployment. Content sprawl has traditionally been associated with files and documents. However, as the usage of more complex Microsoft 365 objects grows, sprawl can include groups, teams, sites, Loop workspaces, Power Apps and Power Automate flows.

When a user is licensed for Copilot, it is embedded across all their major Microsoft 365 apps and can be used to speed up the creation of content, workspaces and apps. While this should facilitate numerous productivity benefits, Gartner believes that, if left unchecked, Copilot will significantly increase the amount of sprawl within Microsoft 365.

If left unchecked, Copilot for Microsoft 365 will usher in a new era of AI-generated sprawl.

One of the main reasons for this is that Copilot will lower the barrier to entry for the creation of self-service but more complex Microsoft 365 objects. If we consider Power Apps and Power Automate, while these are designed to be low-code citizen developer tools, complexity can quickly ramp up. For many users, this complexity might put them off building solutions in these platforms. However, with Copilot, the user will simply describe the automation or app in natural language and the AI will attempt to build it. Figure 8 shows an example of this in Power Automate.

Figure 8: Copilot for Microsoft 365 Creating a Power Automate Flow



Similar shortcuts will exist in Copilot for creating new teams, groups and sites. While reducing complexity and lowering the barrier for entry is undoubtedly a positive development, organizations must be ready for a potential explosion in content, workspaces and apps. Proportionate governance and life cycle policies must be developed to mitigate this new era of sprawl.

Mistakes

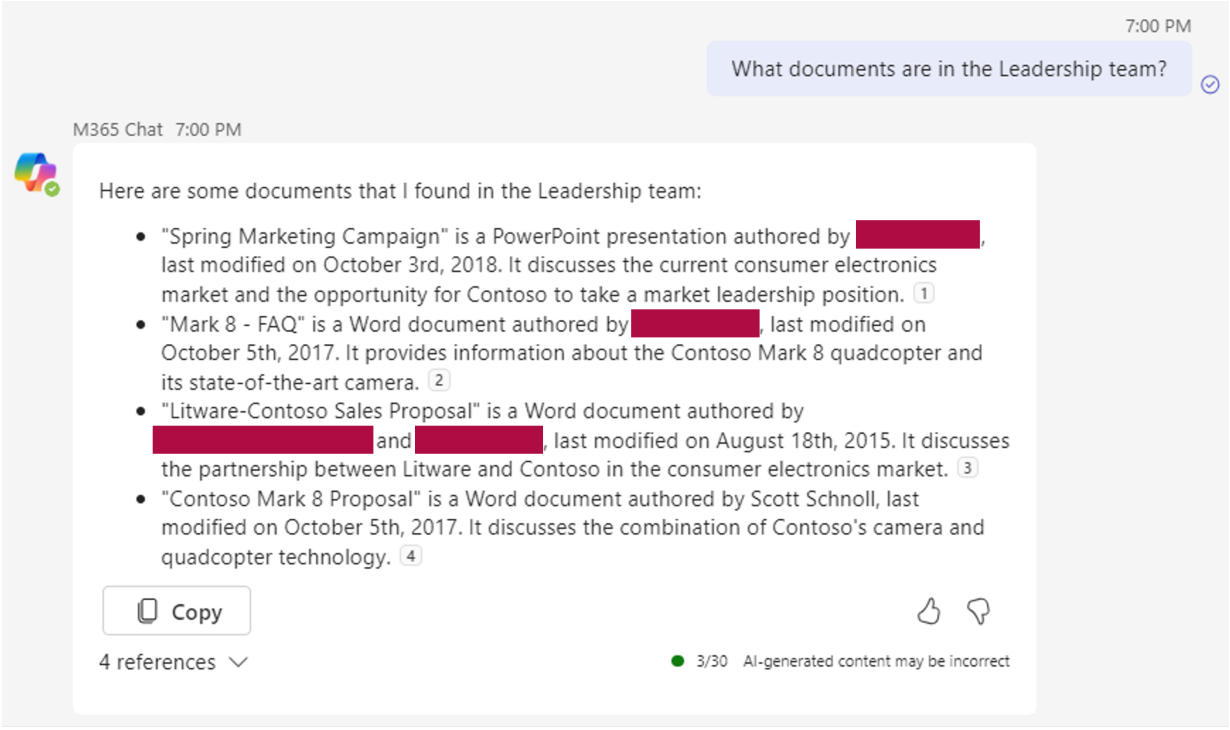
Generative AI can make mistakes, sometimes referred to as “hallucinations.” Even if the user has access to the right information, there is no guarantee that the LLM will provide the right answer. During a 2023 Gartner analysis of social media data across 12,573 risk leaders, 17% of conversations focused on the risk of bias and hallucinations in generative AI. [6] This issue is not specific to Copilot for Microsoft 365. Indeed, many of ChatGPT’s mistakes and biases are well-documented. However, as Copilot has access to your organizational data, these mistakes could be more consequential.

During our testing, we observed some examples where Copilot made mistakes. For example, in our test tenant we had a private Microsoft 365 team called “Leadership.” Using a user account that was not a member of the team, we asked Copilot “*What documents are in the leadership team?*” We expected it to say no documents were available because the user did not have access. However, instead it returned a list of documents that it claimed were in the team but were in fact, located in other teams.

Copilot for Microsoft 365 may appear to retrieve the right information when in reality it has made a mistake

Copilot respects permission boundaries, and all of our testing confirms this. However, Copilot had made a mistake — it had found documents from another team, a public one that the user did have access to, and had incorrectly attributed them to the “leadership team.” Figure 9 shows this (names redacted). Mistakes and hallucinations are unfortunately inherent within all generative AI models and Copilot is not exempt from this.

Figure 9: Copilot for Microsoft 365 Hallucination



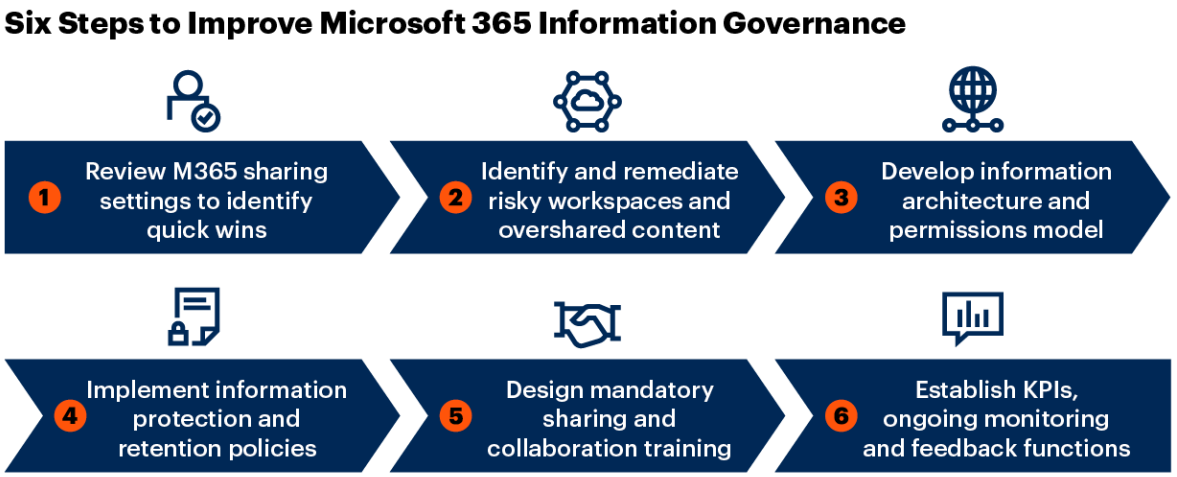
You can reduce the likelihood of user's experiencing generative AI mistakes and hallucinations by providing clear guidance and training on how to write and develop prompts (see the Prompt Training section). In our example, a more precise and contextual prompt may have reduced the chance of Copilot formulating an incorrect response. However, it is still important that users apply skepticism to what they receive back from AI and fully check the sources before relying on this content. In our example, it is easy to see how a concerned user could contact IT or escalate to leadership over a perceived oversharing risk that was not real.

Improve Microsoft 365 Information Governance to Mitigate Risk

Back

To mitigate the above risks, and to enable your organization to get the most benefit out of Copilot for Microsoft 365, you must revisit and improve your Microsoft 365 information governance. Gartner recommends six steps for doing this which are shown in Figure 10 and expanded upon below. Note, these steps do not need to be followed sequentially and it is expected that the longer-term initiatives (such as developing your information architecture, designing your training plan, and implementing data protection and retention), will be run in conjunction with each other.

Figure 10: Six Steps to Improve Microsoft 365 Information Governance



Review Microsoft 365 Sharing Settings to Identify Quick Wins

The default SharePoint sharing settings for Microsoft 365 are very open and can lead users to easily overshare content. Some organizations have locked these down, but Gartner has spoken to many clients that have left gaps.

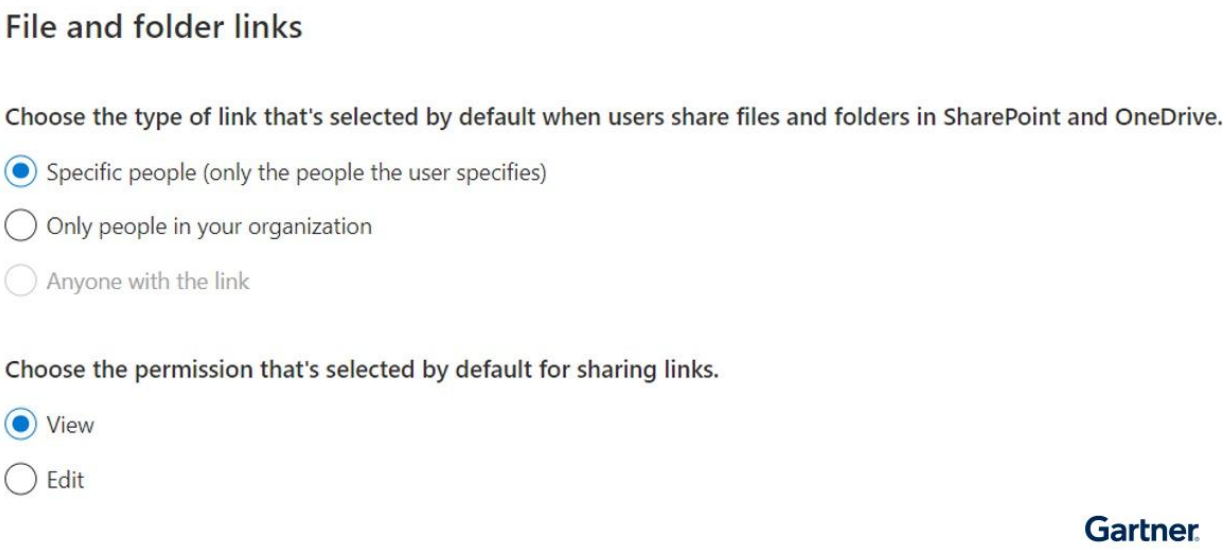
In the “Sharing” section of the SharePoint admin center, Gartner strongly recommends disabling “Anyone” links. This is a form of unauthenticated sharing and users can easily pass files shared using Anyone links to any internal or external user. If you have a business reason to allow these links, you must implement a short expiration period. For more information on the risks of Anyone links and how to implement effective external sharing in Microsoft 365, see Guidance Framework for Managing External Sharing in Microsoft 365.

Under the “File and folder links” section, Gartner recommends changing the default link type to “Specific people” and the default permissions to “View.” A common source of oversharing within an organization is when users select the “Only people in your organization” link to share content from their OneDrive or a SharePoint site. Sharing a file in this way means that every user in your organization has permission to access it. Making this the default option can be dangerous because users don’t have to think about who really needs access. It should be highlighted that files shared using these links are only indexed when they are explicitly sent to a user, or when a user accesses them. As such, an end user should not be able to find content shared in this way using Copilot or search, unless they have previously interacted with it. That being said, this is relying on “security by obfuscation,” and it is very easy for users who have discovered the link to forward it on to others. As such, Gartner recommends using the “Specific people” option by default.

Gartner strongly recommends that you adopt a policy of “just enough access” in Microsoft 365.

By changing the default to “Specific people,” you will encourage users to share content with only those that need access. Figure 11 shows the settings that Gartner recommends.

Figure 11: SharePoint Default Sharing Links



While changing the default sharing link type is an incentive to do the right thing, it does not prevent users from selecting the “People in your organization” link and sharing with the whole company. Users often do this because it is easier than sharing with named individuals or groups. You can disable this link on a site-by-site basis by using PowerShell. In the example below, the code targets a user’s OneDrive site and disables companywide links. Once this takes effect, the user no longer has the option to share using the “People in your organization” link. This could be adapted to apply to multiple sites across both SharePoint and OneDrive. In the longer term, it would require activity auditing/monitoring to cater for new users and sites.

$site = Get-SPOsite https://test-my.sharepoint.com/personal/max_goss_test_onmicrosoft_com

Set-SPOSite -Identity $site DisableCompanyWideSharingLinks Disabled

---

It is also possible to restrict the default sharing link at a site level to “People with existing access.” This could be useful when sharing content from a Microsoft Teams connected site. Members of the team already have access to files that are stored in the team, so defaulting the link sharing type to those with existing access would be a useful step to counter oversharing. Example PowerShell code is included below:

$testTeam = Get-SPOSite https://test.sharepoint.com/sites/max-admin-test-team

Set-SPOSite $testTeam -DefaultLinkToExistingAccess $true

Finally, you can restrict who can share content in SharePoint sites to site owners only. This is useful if you want to restrict members from sharing content with additional users outside of those that already have access to the site. This can be achieved on a site-by-site basis through the SharePoint UI or through PowerShell as shown below:

$testTeam = Get-SPOSite https://test.sharepoint.com/sites/max-admin-test-team

Set-SPOSite $testTeam -DisableSharingForNonOwners

For more information see Limit Sharing in Microsoft 365, Microsoft Learn

Note, while some of these changes can be described as “quick wins,” they must be prefaced by appropriate end-user communications and change management. They should be delivered as part of a communications campaign that aims to promote awareness about how best to store and share content in Microsoft 365 and reduce risk.

Identify and Remediate Risky Workspaces and Overshared Content

In Gartner’s 2023 Microsoft 365 Survey, 42% of respondents noted that they do not measure Microsoft 365 usage/adoption at all.

You can’t govern what you don’t see. Organizations must gain visibility across their Microsoft 365 workspaces and content, and apply appropriate policies to ensure compliance.

Discovery is one of the first steps that organizations need to take to improve their information governance in Microsoft 365. In most cases, organizations are surprised to learn that they have far more overshared content than they thought.

Risky Workspaces

Focus on identifying and remediating access for overshared Microsoft 365 containers or workspaces (groups, teams and sites). Any workspace that is set to “public” or is shared with the “Everyone except external users” group, can be accessed by the entire organization. If any documents are stored in these workspaces, Copilot and search can find them. Gartner has seen examples where organizations have paused Copilot pilots because the tool was returning sensitive information that had been stored in the wrong location.

To mitigate these risks, first, perform an audit of your Microsoft 365 workspaces. Focus on groups, teams and sites that are either public or are shared with a large number of users or groups. For example, if you have an “All-Staff” security group, check to see what SharePoint sites are using this group and if the content is actually appropriate to every member of staff. Even if you have a workspace provisioning process, remember that workspace owners can change the privacy settings of their workspaces. Contact the owners of public workspaces and check whether these spaces and their content should actually be available to everyone in the organization. This is particularly important in large organizations with different user populations, for example, a university with staff and students.

Note: as a short-term measure, you could consider excluding certain high-value risky workspaces from search results. You can do this by accessing the relevant SharePoint site as a site administrator. Click the gear cog, select “Site Settings” > “Search and offline availability,” and select “No” under “Allow this site to appear in search results” This will remove the site from the search index and prevent search or Copilot from finding any documents or information in the site. **This should be used as a short-term step for highly sensitive sites only**. Hiding a site from search does not address the underlying permission issues and could negatively impact the user experience.

The below PowerShell example demonstrates how to return a list of Microsoft 365 groups/teams where the AccessType is “Public.” This uses Exchange Online PowerShell.

Connect-ExchangeOnline -UserPrincipalName #Enter admin username

$M365Groups = Get-UnifiedGroup | Where-Object {$_.AccessType -eq “Public”} | select “Name”, “DisplayName”, “Description”, “ResourceProvisioningOptions”,

“SensitivityLabel”, “AllowAddGuests”, “accessType”, “GroupMemberCount”, “GroupExternalMemberCount”, “ManagedBy”, “WhenCreated”, “WhenChanged”

The script can be downloaded in full here:

Get Public Groups Script

Identifying “Public” groups and teams is a good starting point. To further reduce risk, expand your search to consider workspaces with the following characteristics:

- “Public” groups or teams, or those shared with the “Everyone but external users” group
- SharePoint sites with an incorrectly used “All staff” group
- Ownerless groups, teams or sites
- Groups, teams or sites with no recent activity
- Groups, teams or sites with large numbers of All company/Anyone links
- Groups, teams or sites with lots of external guest users
- Very large SharePoint sites (more than 100GB of stored data).

- Workspaces with a “public” or “unrestricted” sensitivity label that contains sensitive or PII information.

Once you have performed an initial audit and cleanup, Gartner recommends developing a workspace inventory for your organization. This should contain key metadata to determine: what the workspace is for, which department it belongs to, who owns it and whether it contains sensitive or personal identifiable information (PII) data. As part of this, you should leverage container-level sensitivity labels (see the Implement Information Protection and Retention Policies section), to help classify sensitive containers and apply relevant rules. Examples include, restricting users with unmanaged devices to web-only access with no download, or preventing external sharing from the workspace.

Use the workspace inventory to define what makes a “compliant workspace” and set up automated policies to remediate workspaces that fall out of compliance. As an example, an organization might define a compliant Microsoft team as having:

- A minimum of two owners
- A consistent naming convention
- Access controls set to “private”
- Completed metadata that defines the team
- Compliance with a specified storage quota
- No explicit permissions outside of team members and owners.

If something changes in the team that makes it “not compliant,” for example, a file is shared explicitly, this is caught by automatic policies that periodically run and take action to bring the team back into compliance. This may involve contacting the owner, and/or applying a remedial action (such as removing the explicit sharing link or switching the team back to “private” if it had been changed to “public”).

To combat future sprawl, combine reactive policies with proactive methods that enforce compliance when a workspace is first created.

Gartner recommends creating a managed self-service process to govern the creation of new teams, groups and sites. This is usually achieved through a custom Teams app built using Microsoft Power Apps, SharePoint Framework (SPFx), or a third-party solution. When coupled with reactive policies that remediate noncompliant workspaces, this mechanism is highly effective at reducing the risk of over shared workspaces in Microsoft 365 (see Assessing Workspace Governance and Life Cycle Controls in Microsoft 365).

Overshared Content

Identifying and classifying overshared content that Copilot will use in a response is challenging. Gartner recommends that you start by looking at content that has been shared with the whole organization, or through "Anyone" links (if enabled). These links, along with security groups such as "Everyone except external users," can pose an oversharing risk if they have been used to share sensitive content. While It can be difficult to audit these links and permission groups, it is important in terms of understanding the patterns of sharing used within your organization.

Prepare your information for Copilot for Microsoft 365, by preparing it for search

Before assigning Copilot licenses, perform tests in Microsoft Search using nonprivileged accounts. If you find content that you should not have access to, this is an indication that you should do further remedial work before deploying Copilot. You can also test for oversharing by using the Delve profiles. As an example, if you access an executives' profile and look at the files tab, is there anything in here that you should not see? These spot checks will not be conclusive, but can be a useful first step as you test for overshared content.

To get a more holistic view, organizations with Microsoft 365 E5 can use the file reports within Windows Defender that can scan files across your Microsoft 365 environment. You can define filters to show content shared across the whole organization, with specific security groups, or with external users. If you are using Microsoft Purview's Information Protection, you can also filter by files that have been tagged with sensitivity labels, which should offer further insights into whether sensitive content is being overshared (see File filters — Microsoft Defender for Cloud Apps, Microsoft Learn).

If you have Microsoft 365 E3, then you can get similar data from the Microsoft 365 audit logs, although you will need to develop a process to extract it. Note that audit logs for E3 organizations are kept for 180 days, for E5 organizations they are retained for 12 months. You can target specific sharing and access requests, for example “Created a company sharing link” or, “Created an anonymous link.” These can be cross-referenced with data from Microsoft Graph to determine which areas in the organization are relying heavily on these sharing methods (see Use Sharing Auditing in the Audit Log, Microsoft Learn).

You can also leverage the Information Oversharing template, which is part of the Microsoft Graph Data Connect solutions. Here you can build a report utilizing an Azure data store to analyze sharing information across your SharePoint estate at scale (see Information Oversharing Template Setup, GitHub).

Note that third-party products are often required to identify and manage containers and content at scale in Microsoft 365. These are discussed in the Assess Add-On Products to Further Improve Visibility and Governance section.

Develop Information Architecture and Permissions Model

How to design and build a future proofed Information Architecture across Microsoft 365 storage locations is a common question among Gartner clients. SharePoint is not a like-for-like replacement for an on-premises file server, and it does not work well with a rigid hierarchical structure. Consider this step as part of your longer-term strategy to improve your information governance in Microsoft 365.

Develop a Flat Architecture to Distinguish Between Self-Service and Authoritative Content Sources

To minimize the complexity of nested permissions, Gartner recommends creating a flat architecture where site collections, document libraries and shared channels are used to define permissions. Avoid, where possible, creating unique permissions at the folder level or using subsites. These add to complexity and can cause confusion for users in terms of both where to store content, and who has access. This, in turn, can lead to oversharing, which can be easily exposed through a Copilot response. To prevent users creating new subsites and introducing more complexity, Gartner recommends disabling subsite creation in the SharePoint admin center.

Sites should be grouped together with hub sites that can provide consistent branding, navigation and search across all connected sites. For example, an IT hub might serve as a landing point for the whole division with sub departments or related business processes having their own SharePoint site. Avoid recreating a shared drive folder structure in SharePoint. Rather than storing millions of documents in a single library, split your content out across different sites and libraries. Doing so will improve search and performance, and reduce the amount of nested permissions.

While OneDrive, Teams and SharePoint all use a SharePoint site as their back-end storage, it is important to highlight that they have very different roles:

- **OneDrive:** Is used to store individual content and drafts. Business-critical records or published organizational content should not be stored in OneDrive.
- **Microsoft Teams:** Is well-suited to storing "working" documents. A good use case is where a project or team has created a team to store content that is in development or is related to a particular physical team. It is not well-suited for storing organizational content or for holding business-critical records.
- **Managed SharePoint sites:** These are not connected to a Microsoft team and are typically provisioned by IT. They are used to hold departmental or organizational content or to support a key business process. They typically have many consumers (readers) but a small number of editors. These sites hold finished policies or procedures and business records, and may contain content that is relevant for a larger audience.

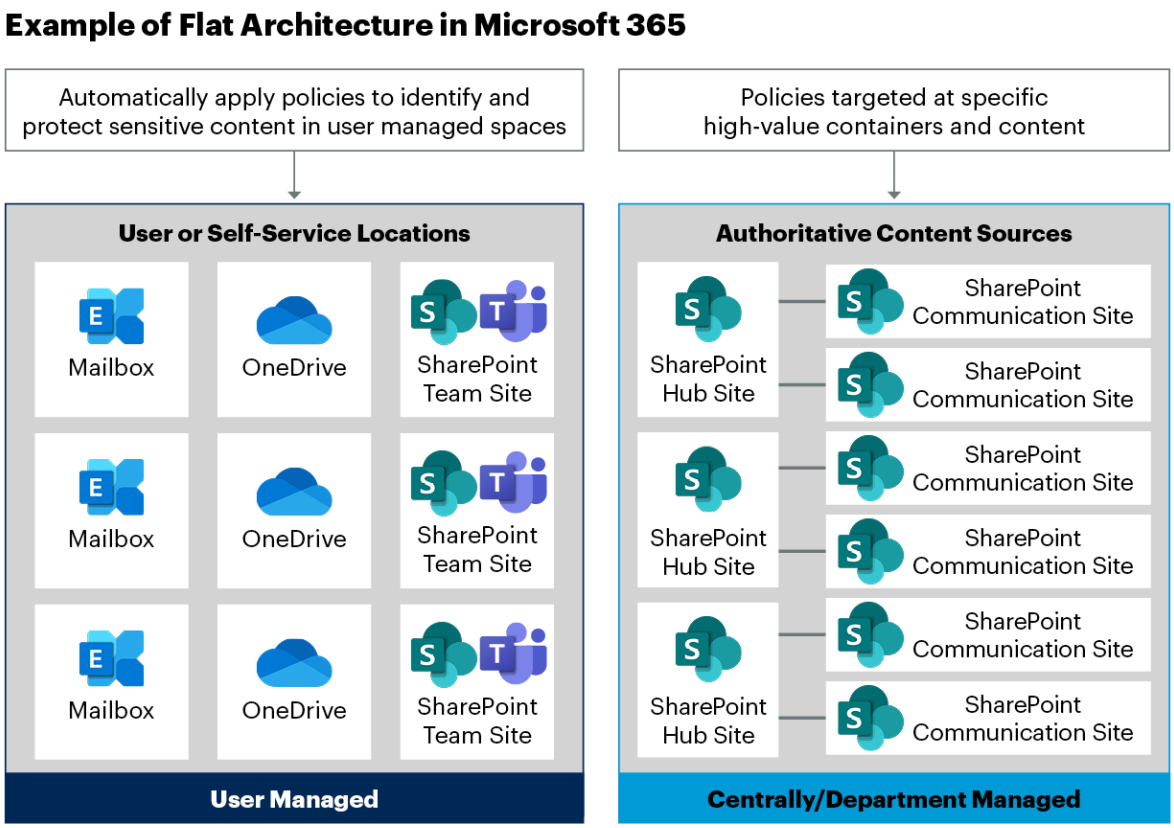
User-Managed Versus Centrally Managed Content Sources

There are two types of content sources in Microsoft 365 — one that belongs to the user (OneDrive, user mailbox), or is managed by the user for a group of users (Microsoft team or shared mailbox) and the second, that is created at the department or organizational level and is designed to be centrally managed and to hold authoritative content. Managed SharePoint sites fall into this second category. They typically hold finalized content (for example, policies and procedures) and are available to a wider audience.

Figure 12 shows an example of a flat architecture in Microsoft 365. On the left we see containers that are user owned or managed. On the right, are containers that are managed by a department, or on behalf of a significant business process. These spaces are more structured and are essentially known authoritative content sources. For example, if a project has created a new HR policy, it should be stored in the appropriate HR-managed SharePoint site, rather than being left on a team, or in a user's OneDrive.

As shown in Figure 12, organizations should apply policies that go across their Microsoft 365 content estate so that, if sensitive information is incorrectly stored, it can be protected wherever it is. However, by having defined locations to store business-critical or department-level content, organizations can improve how they manage content, and crucially how users find and share it. They can also be sure that they are protecting and retaining their most important information by applying policies and controls to these locations.

Figure 12: Example of a Flat Architecture in Microsoft 365



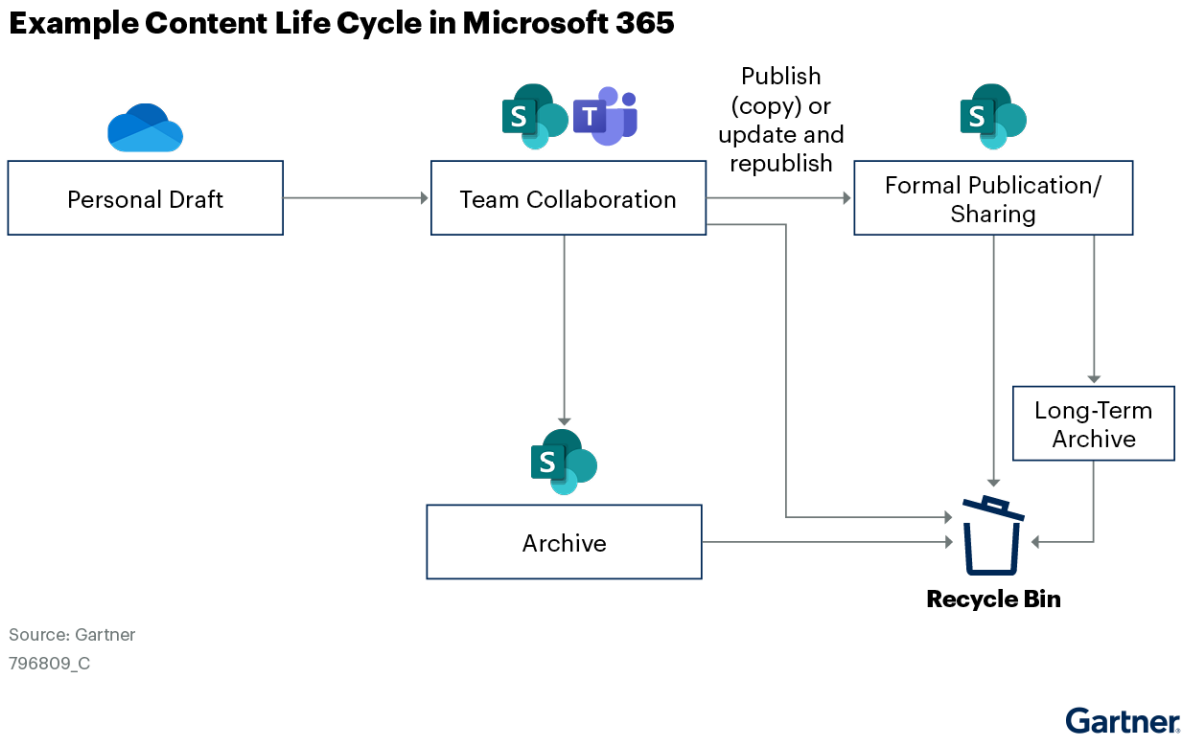
Source: Gartner
800669_C

Gartner

Importance of Content Life Cycle

Figure 13 shows an example content life cycle in Microsoft 365. Content might begin as an idea in a user's OneDrive. It is then moved into a Microsoft team as it becomes part of a project or initiative that multiple users must work on before being published to a managed SharePoint site for wider consumption. Critically, redundant content and containers (including groups, teams and sites) should be archived or deleted to reduce storage consumption and mitigate the risk of users finding and sharing outdated information.

Figure 13: Example Content Life Cycle in Microsoft 365



Gartner has spoken to many clients whose users have overshared because their content is stored in the wrong place. As an example: Consider a finalized policy that should be accessible to the whole of IT, but is stored in a Microsoft team, where only the users in that team have access. Tiring of fielding access requests, the team's owner decides to share the policy with the whole organization using a companywide link, or copies the file to a public team. Both approaches result in oversharing as the policy, which was meant to be available only to IT, is now available to the whole organization. Now this policy can be easily exposed to the wrong audience through Copilot and search.

Permissions Management

Building a well-defined permissions model with clear directives for users on where to store content is fundamental to reducing oversharing risks in Microsoft 365. Central to this is automating permissions management as much as possible. When users are required to spend too much time manually managing permissions, oversharing will often occur.

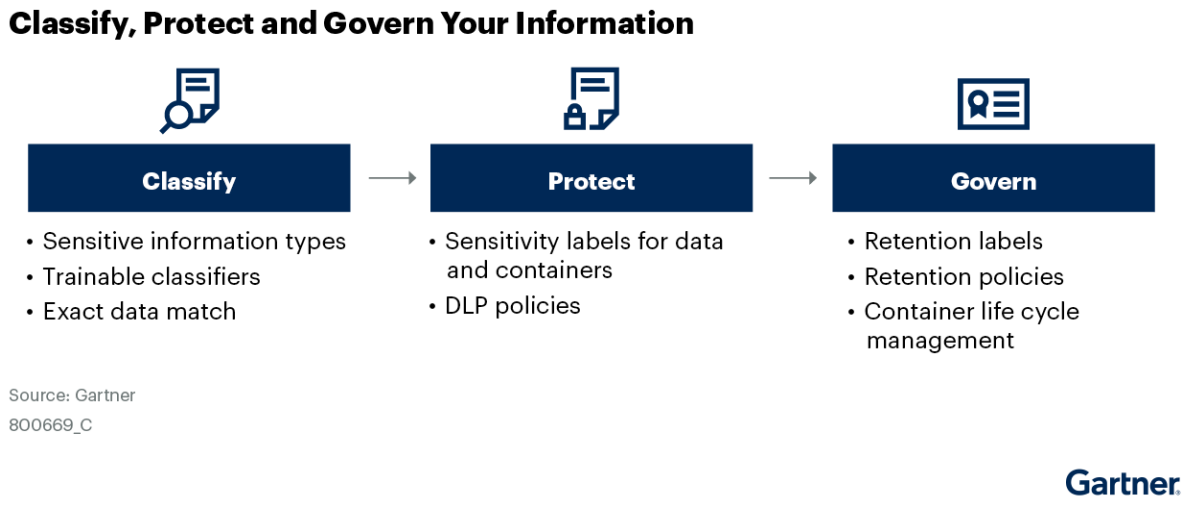
Gartner strongly recommends that you create a series of organizational groups that are mapped to your companies' structure and are sourced from your identity management system. These groups (commonly mail-enabled security groups), can be used to dynamically drive permissions in your managed SharePoint sites. As users join or leave the organization, or change departments, group membership automatically updates and your site membership is kept accurate. For example, you may create an HR authoritative group and apply it to the HR hub site and connected managed SharePoint sites, using the group to drive view permissions. The messaging to end users is now very simple, if you need to share content with the whole of HR, you can store the content on one of these sites, and anyone in HR automatically has access. You do not need to worry about managing permissions, or creating companywide sharing links.

For user-managed spaces, a dynamic permissions group may not be appropriate as the group or team owner should be able to invite users to the workspaces. Organizations with the Microsoft Entra ID P2 (formerly Azure AD P2) license should use access reviews to prompt owners to check the members of their groups and teams and determine if they still need access. Access reviews can be used to review both internal and external members and can contain a default action if the owner does not respond (such as remove access). Similar functionality could be achieved through PowerShell or third-party management tools. Note, to stay on top of permissions management, this must be an ongoing activity.

Implement Information Protection and Retention Policies

To fully protect and govern your information and mitigate risk when deploying Copilot, you need to use a combination of Microsoft 365 security and compliance solutions. Although organizations with an E3 license have access to *some* of these features, to take advantage of the automatic scanning and application of sensitivity and retention labels, E5 or an appropriate E5 add-on is required. The full implementation of Microsoft Purview solutions (such as Information Protection, Data loss Prevention (DLP) and Records Management) are complex subjects and beyond the scope of this research. The following sections provide an overview of how these services can be used to help classify, protect and govern sensitive information. Figure 14 outlines the approach, with more detail provided below.

Figure 14: Classify, Protect and Govern Your Information



Information protection and retention are two sides of the same coin. Together they can be used to apply appropriate policies to protect and govern content throughout its life cycle.

Classify and Protect Sensitive Information

Before you can protect sensitive information, you first need to be able to identify and classify it. This can be achieved through a combination of Microsoft Purview solutions including Sensitive Information Types (SITs), Exact Data Match (EDM) and Trainable Classifiers. These are all located in the Data Classification section of the Microsoft Purview admin center. SITs use pattern matching to automatically detect whether an item contains sensitive information. Microsoft provides several preconfigured SITs — such as credit card and bank account numbers — or organizations can create their own. EDM goes further by allowing organizations to define exact data values rather than generic patterns. Trainable classifiers enable an administrator to train Microsoft 365 to recognize certain types of content and automatically apply sensitivity labels. This feature could be used, for example, to recognize a legal agreement and automatically apply a sensitivity label that restricts access to a specific group of users. Once you have classified your content and defined what constitutes sensitive information for your organization, you can use a combination of DLP and sensitivity labels to protect it (see Microsoft Purview Information Protection).

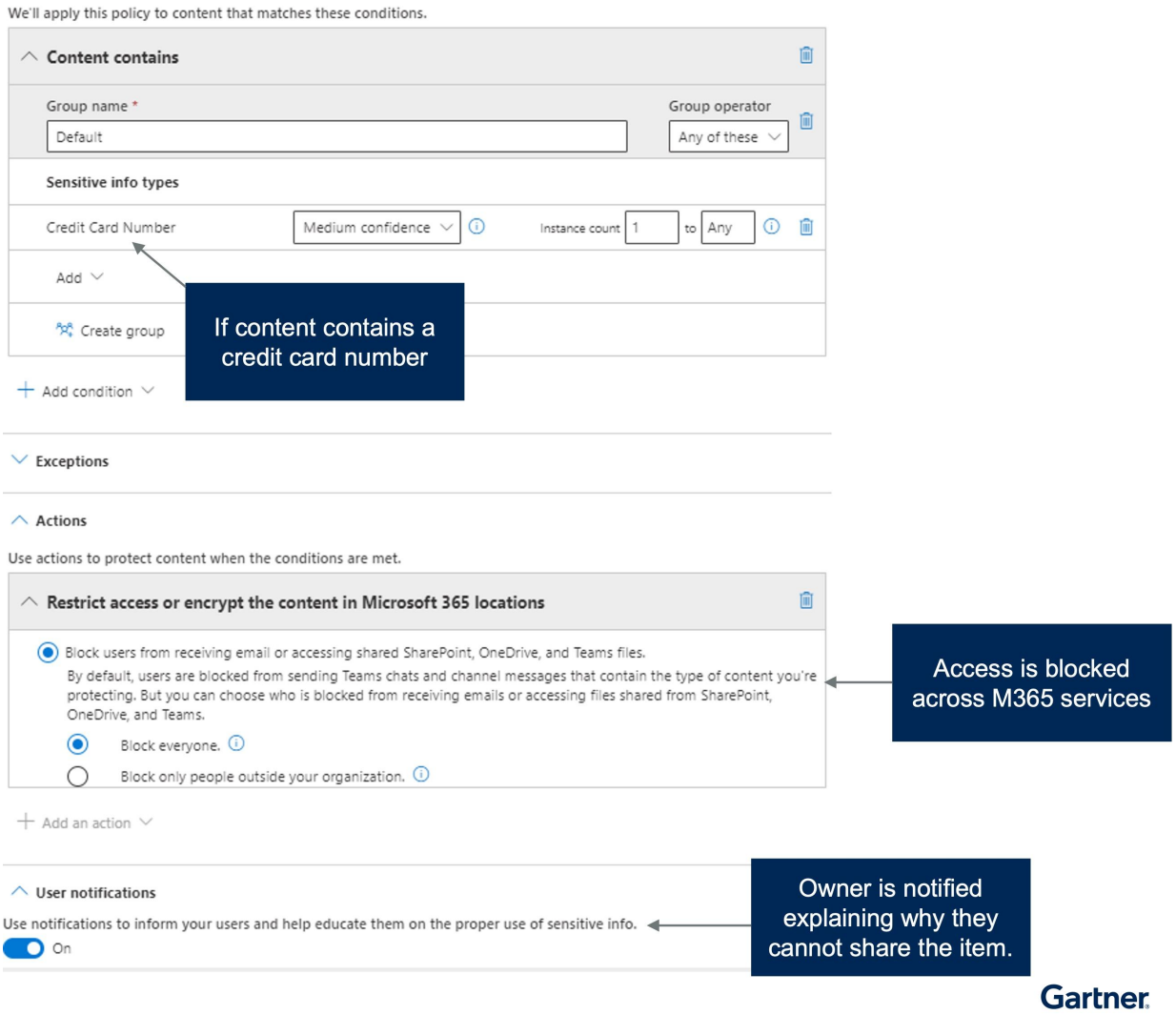
Data Loss Prevention Policies

You can use DLP policies to identify, monitor and automatically protect sensitive information, preventing oversharing and reducing the chance that it will be accessible to Copilot or search. DLP in Exchange, SharePoint and OneDrive is available to organizations with E3. DLP for Teams chat is only available with E5.

DLP policies are flexible and can be configured to warn the user rather than fully block. They can also notify an admin if a user action triggers the policy, (such as a user trying to share restricted content). Both sensitivity labels and classifiers such as sensitive information types can be used as conditions for a DLP policy.

Figure 14 shows an example of a DLP policy. In this case, we are using one of the built-in sensitive information types which detects credit card numbers as an input to the policy. If a credit card number is detected in a document stored in SharePoint or OneDrive, all sharing is blocked and the content owner sees a customizable notification explaining why they cannot share the document. DLP can be used effectively to prevent the oversharing of sensitive information and can also be used as an education tool.

Figure 14: Example DLP Rule



Sensitivity Labels

Sensitivity labels are another important component of your information protection strategy. They can be applied to both individual files and to containers (such as groups, teams and sites). When applied at the file level, labels can add visual markings (watermarks) and/or add encryption, which prevents unauthorized access. They also serve as a user education tool by visibly labeling the file (such as Unrestricted, General, Confidential or Highly Confidential) and providing a description about how the file should be used. For example, files labeled as “Confidential” should only be shared with authorized named individuals. Labels travel with the file, even if they are moved from their original location. With E3, administrators can publish sensitivity labels to users who can manually apply them to their files. With E5, labels can be applied automatically based on specific text or patterns within the file (such as a credit card number), or based on file properties or metadata.

Copilot for Microsoft 365 respects all access controls enforced through sensitivity labels and DLP policies

Consider a scenario where a sensitivity label is applied to a document that encrypts it and restricts access to a specific security group. Users who are outside of this group (unauthorized users) cannot use Copilot to find, synthesize or create information from that document, even if it has been shared with them. When a sensitivity label applies encryption, the user must have both the VIEW and EXTRACT usage permissions for Copilot to be able to find or interact with that content.

When an authorized user instructs Copilot to generate new content from a document that has a sensitivity label applied, the new content will inherit the original label. If Copilot synthesizes data from multiple sources with different sensitivity labels, the label with the highest priority (which should equal the most restrictive) is applied.This also applies to Copilot’s chat interface where any response that references content with a sensitivity label will inherit that label. By ensuring that your sensitive content is appropriately labeled, you will help to ensure that it stays protected when Copilot interacts with it. Microsoft provides more information on how sensitivity labels and other Microsoft Purview controls can be used with Copilot for Microsoft 365 here: Protect and manage Microsoft Copilot for Microsoft 365 by using Microsoft Purview.

Developing a clear labeling strategy is essential to being able to identify and apply suitable protection to your content over the long term. For example, content classified as "unrestricted" is likely to be suitable for sharing with a wider internal (and potentially external) audience. However, content classified as "confidential" will likely need to be locked down to authorized individuals. Adopting a labeling strategy will help you to protect sensitive data while empowering users to share content that is suitable for wider consumption. Gartner recommends using sensitivity labels at both the content and the container level (groups, teams and sites). More information on sensitivity labels is available in Gartner's Solution Path for Document Management in Microsoft 365.

Govern Information With Retention and Records Management Policies

While organizations rightly focus their efforts on protecting sensitive business information, data life cycle policies and the appropriate deletion of redundant or stale content is often overlooked. Oversharing is not the only information governance risk that Copilot for Microsoft 365 accentuates. Misinformation, which often stems from users accessing inaccurate or stale content, is a real risk for organizations, and one that will likely increase in the generative AI content age.

If Copilot does result in an explosion of AI-generated content, containers and apps, it will become even more important to utilize retention and deletion policies to manage your content estate. As well as developing targeted policies to retain business-critical information, organizations should develop deletion policies to remove redundant or stale content.

Microsoft 365 E3 offers some retention capabilities. Namely, you can use broad retention policies that retain content across entire SharePoint sites for a specific period of time. You can also publish labels to specific locations and encourage users to manually tag files they wish to retain/delete, but this often results in inconsistent adoption. E5 (or the E5 compliance add on) offers far more functionality, including the ability to automatically apply retention labels at scale, and to use records management features including file plans, disposition reviews and defensible destruction. For more information on the differences between E3 and E5 from a retention and records management perspective, see Assessing Records Management Capabilities in Microsoft 365.

Irrespective of whether you use E3, E5, or have a third-party tool, you must avoid applying blanket retention policies to all Microsoft 365 locations.

Retaining all content, irrespective of value, is a common mistake that organizations make when applying retention in Microsoft 365. Blanket policies — for example, retaining content in all SharePoint sites for seven years — results in redundant content being kept alongside important business information. This results in three issues:

- **Organizations run out of SharePoint storage space:** This is a common theme Gartner has observed over the last 12 months. As more content is stored in Microsoft 365, more care must be taken to manage it and ensure timely deletion. Organizations that exceed their SharePoint storage quota usually have to purchase more from Microsoft — and this is expensive.
- **Redundant or stale content appears in search:** Retaining everything — irrespective of its importance — also has the knock-on impact of polluting your search and creating user confusion. As previously demonstrated, redundant, trivial or obsolete (ROT) content could result in Copilot for Microsoft 365 returning irrelevant or outdated information, leading to misinformation.
- **Legal and compliance risks:** For regulated organizations, certain types of information must be disposed of after a set period of time. Keeping it for longer could result in compliance risks and make discovery requests more complicated.

To avoid retaining all content, Gartner recommends targeting labels and policies to high-value locations that you know contain business-critical information. These are likely to be the authoritative content sources (such as managed SharePoint sites). Organizations with E5 licenses should use adaptive scopes and rules-based policies to further refine their scope, targeting specific libraries and content types. For the content in the user-managed locations, avoid blanket policies, but instead look for sensitive information. For example, teams and groups should be subject to an expiration policy, but if a particular team contains business-critical information, these should be tagged and retained (Figure 13 depicts this model). E5 customers should use automatic label policies to label sensitive content based on whether it matches defined sensitive information types or trainable classifiers. E3 customers will likely need to augment Microsoft 365 with a third-party or use the E5 compliance license to achieve similar results.

Retention and Deletion Policies for Copilot Prompts and Responses

User prompts to Copilot for Microsoft 365 — and Copilot's responses back to the user — can be included in Microsoft Purview retention and deletion policies. Microsoft has expanded its Teams chat retention location to include Copilot messages, renaming it to "Teams chats and Copilot interactions." This includes messages from individual chats, group and meeting chats, as well as bot and Copilot interactions. It does not include Teams channel messages, which are handled separately. If you already have a policy setup for Teams chat, this will automatically apply to all Copilot messages. Behind the scenes, Copilot messages work in the same way as Teams chat messages, and are stored in a hidden folder in the exchange mailbox of the user who runs Copilot. Like Teams chat, these messages can be found in an e-discovery search. If a policy is applied that deletes Teams chat and Copilot messages, they will be deleted from the UI first, then deleted from the hidden folder.

Many organizations have set up a relatively short retention/deletion policy for Teams chat (such as 90 days). If you have not already set up retention policies for Teams chat, Gartner strongly recommends that you consider this before implementing Copilot for Microsoft 365 at scale. If you do nothing, these messages will exist permanently and could significantly complicate e-discovery searches as well as lead to future security and compliance issues. They also consume mailbox storage space. Microsoft has more information on retention and deletion policies for Copilot prompts here: Learn about retention for Copilot for Microsoft 365.

Design Mandatory Sharing and Collaboration Training

Most organizations invest heavily in creating mandatory training for security and compliance, but often neglect sharing and collaboration best practices. Just as users should be taught how to avoid a phishing link, they should be shown how to appropriately store and share content in their primary collaboration systems and how to manage sensitive information. Oversharing sensitive information or storing it in the wrong place can lead to significant business risk and, as highlighted in this research, using Copilot will likely accentuate this risk if action is not taken.

Users should be taught sharing best practices such as sharing with named individuals or groups, instead of relying on all-company links. They must also be made aware of the benefits of storing shared content in collaborative workspace where the permissions are already set. For example, sharing content in a Microsoft team, where the team members already have access, is far easier than storing the content in OneDrive and working out who should have permissions. It also means that the content won't be deleted if the user leaves the organization.

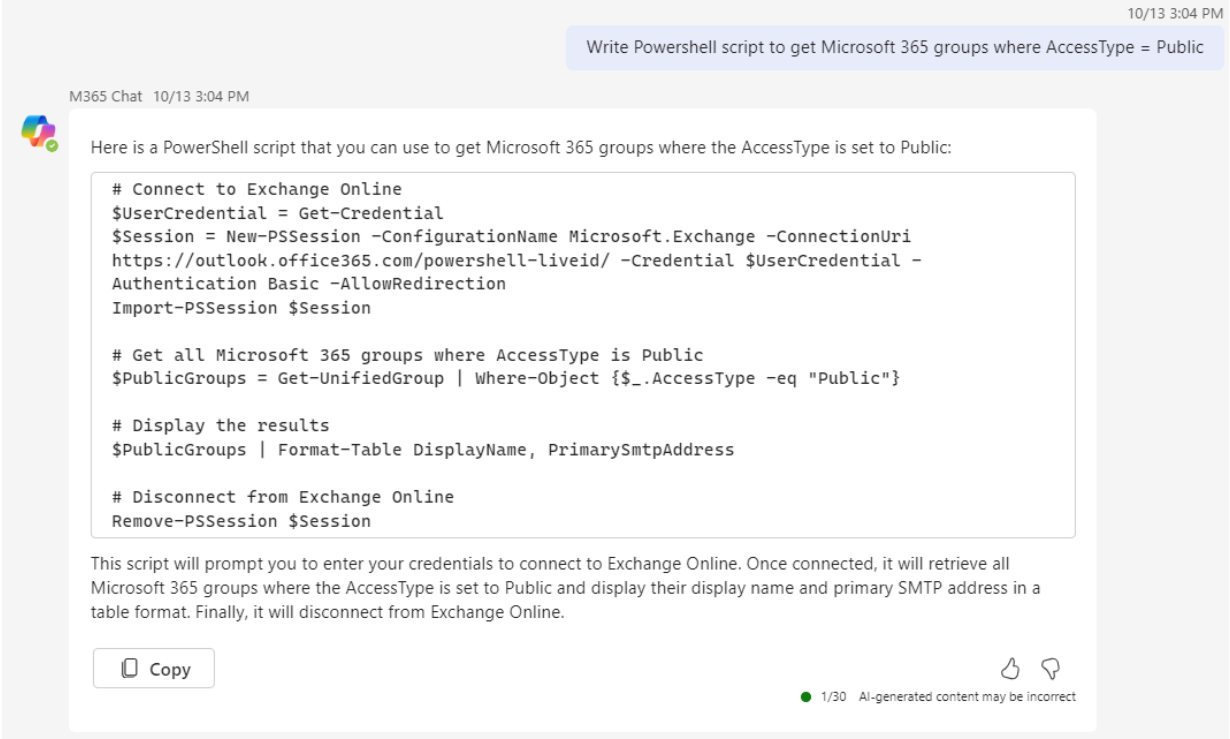
While many organizations offer some of these courses “on demand,” uptake is often low and there is usually little incentive to attend. Instead, Gartner recommends working with your compliance and security leadership to tie these to your mandatory training program, positioning them alongside existing security and compliance modules. If your organization is planning to adopt Copilot for Microsoft 365, use this as a springboard to push for further education and training for end users.

Prompt Training

As organizations embrace generative AI, and use it to improve productivity, user training on these technologies will become essential. The ability to write accurate and contextual prompts, — known as “prompt crafting” or “prompt engineering,” will become a key skill that employees must learn.

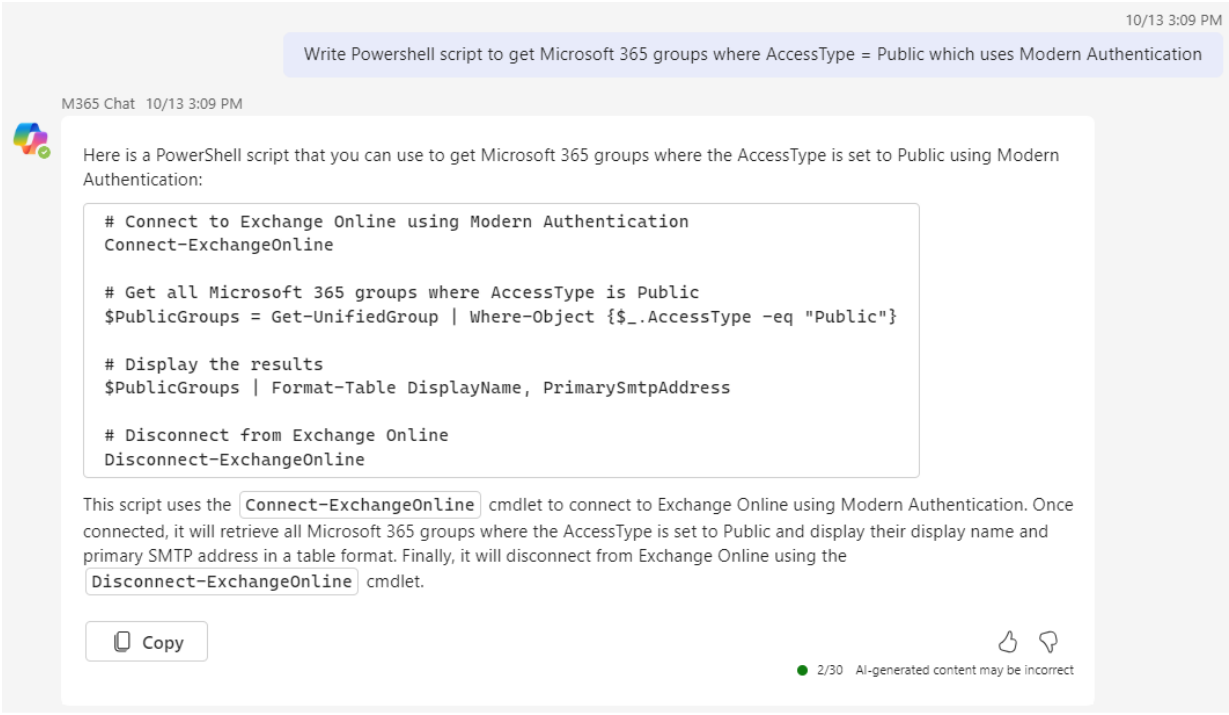
The more precise and more contextual the user prompt, the more likely the response from the LLM will be accurate and usable. Figures 15 and 16 demonstrate this. In Figure 16, we asked Copilot to *“Write a PowerShell script to get Microsoft 365 groups where AccessType = Public*.” Copilot responded with a script that did this. However, the script connected to Microsoft 365 using the legacy version of Exchange Remote PowerShell that Microsoft has deprecated, and is incompatible with modern security standards such as multifactor authentication (MFA). This script could not be used against product environments.

Figure 15: Copilot Prompt Example — Minimal Context



To improve the response, we expanded the prompt providing additional context "*Write a PowerShell script to get Microsoft 365 groups where AccessType = Public* **which uses Modern Authentication**." Copilot's response is shown in Figure 16, a usable script that connects using a supported command that works with MFA.

Figure 16: Copilot Prompt Example — Expanded Context



Gartner

To get the most out of these systems and to reduce the risks of misinformation and mistakes, users must be trained on how to craft accurate prompts, and must have enough domain knowledge to understand when Copilot (or other generative AI tools) are giving incorrect or inaccurate responses. To assist Copilot users, Microsoft provides a Copilot Lab function, which is integrated into Copilot endpoints and provides a centralized platform to help users improve and share their prompts.

Establish KPIs, Ongoing Monitoring and Feedback Functions

To mature your Microsoft 365 information governance, you must develop ongoing reporting, monitoring and remediation capabilities. These will help to ensure that workspaces and content remain compliant. As you develop your Microsoft 365 environment, expand these policies to cover other Microsoft 365 objects such as Power Apps and Power Automate flows. In this context, policies can help to highlight unowned or redundant apps, or personal productivity apps that have become business critical and may need additional governance.

If your organization adopts Copilot for Microsoft 365, you should extend these policies to Copilot prompts and responses. Copilot interactions are captured in the Microsoft 365 audit log and can be found in an e-discovery search. Additionally, you can use Microsoft Purview Communication Compliance to monitor, capture and, if necessary, act on risky Copilot interactions.

In addition to monitoring and remediation, you need to understand how the tools are being used, and where additional training is required. This is particularly important given the surge in AI-generated content, containers and other Microsoft 365 objects that will happen if Copilot for Microsoft 365 adoption is widespread.

In Gartner's 2023 Microsoft 365 survey, only 8% of respondents noted that they were using established KPIs to measure how users were using the platform. [4]

Many organizations are not setting up KPIs or other measurable objectives to define "what good looks like." For example, organizations may be seeing an increasing number of teams or SharePoint sites in their environment, but this is of little value if they do not know what this means. Instead, organizations need to understand what they are trying to achieve strategically from their Microsoft 365 deployment, and build KPIs to measure how the tools are being used. Is there a correlation between more SharePoint usage and less emails being sent? Are they hoping to improve collaboration by having more link-based sharing and less emails with file attachments. Are they hoping to reduce reliance on legacy storage solutions? Defining what good looks like from a collaboration standpoint, and setting up key metrics to help you determine if you are meeting your objectives, will help you to assess how Microsoft 365 is landing across your organization. Crucially, this helps to determine where you should target your communications and training.

When establishing your KPIs, you should investigate how to measure Copilot usage and value. This will be a key requirement for any Copilot pilots or early-adoption programs you are running. Has Copilot resulted in a tangible time saving/productivity gain. For example, its ability to quickly find content and provide summarization has saved X time per Copilot user per day. For those planning Copilot pilots, Gartner recommends a hands-on approach where you regularly engage with your pilot users to understand how they are using the tool and where they are getting value. Establish a clear review point (for example after 90 days), and survey your pilot users to solicit their opinions. For example, what percentage of users are still using Copilot daily, how much time has it saved them, are they able to use Microsoft 365 apps more effectively?

To augment this qualitative data, you should utilize reporting and monitoring tools. Copilot monitoring tools are in their infancy, but as previously mentioned, you can search for Copilot interactions in the audit logs. These are listed as “Interacted with Copilot.” Audit log events include how and when users have interacted with Copilot, what Microsoft 365 apps they were using, and what files (if any) were accessed including their sensitivity labels (if applied). Microsoft also has a basic Copilot usage report available in the reports section in the Microsoft 365 admin center.

Gartner recommends using a combination of reporting tools and policy engines to assess and act on quantitative data, together with feedback functions that give you qualitative insight into how Microsoft 365 tools are being used. Develop communities of practice with engaged end users and run monthly meetings to give you the opportunity to both educate and listen to your users. Utilizing Microsoft Viva Engage (formerly Yammer) to build an online community with a community manager to solicit feedback and encourage engagement, is also a useful way to understand how users are engaging with Microsoft 365 and where the pain points are.

Assess Add-On Products to Further Improve Visibility and Governance

While Microsoft 365 has a wealth of analytics and governance features, these are split out across multiple admin centers and there are gaps in the native controls. Many organizations resort to complex PowerShell scripts to get the required visibility and governance over their Microsoft 365 environments. However, these can be difficult and costly to maintain. Complex and highly regulated organizations should consider add-on products to improve their Microsoft 365 information governance.

Third-Party Add-On Products

Microsoft 365 has an expansive third-party ecosystem, and there are multiple vendors that offer mature products to improve governance, visibility and access controls. Examples include AvePoint, SysKit, Orchestry, Powell Software and Rencore. These products enhance visibility and offer ready-made solutions for common governance requirements, such as insights into public groups, overshared content, and risky sharing links. Many also offer provisioning systems to allow users to create new groups, teams and sites while enforcing appropriate guardrails. One of the key advantages of these products is that they contain policy engines that administrators can configure to automate remediation and life cycle tasks. An example includes a policy to automatically bring noncompliant workspaces back into compliance, or to take action if PII data is found in a workspace that has been classified as “public” or “unrestricted.” One of the key points made in this research is the need for organizations to get their information ready for Copilot. The discovery, insight and automation features offered in these tools can make this task significantly easier.

First-Party Add-On Products

In 2022, Microsoft released its own add-on product designed to improve the governance and management of SharePoint and OneDrive. Microsoft Syntex — SharePoint Advanced Management (SAM) is an E3/E5 add-on license, available at a list price of $3 per user per month. SAM offers extended insights in the form of data access governance reports that allow admins to see the sites that contain the most amount of all company, anonymous, or externally shared links. It also enables administrators to lock down SharePoint site access to a specific security group. Users outside of this group have no access to the SharePoint site or its content, even if it was already shared with them. SAM also allows you to lock down OneDrive access to a specific security group. It is good to see that Microsoft now offers a first-party solution that can assist with governance and oversharing, and that SAM addresses some of the most-common pain points in SharePoint management. However, SAM does not yet have the breadth of functionality to compete with the more mature third-party offerings from an overall platform visibility and governance perspective.

Recommendations

Back

- **Establish a clear information architecture, permission, and retention model for Microsoft 365:** Focus on OneDrive, Teams, and SharePoint, what information they should store, who should have access, and what life cycle rules apply. Build a flat information architecture to avoid complex, nested permissions and adopt a “just enough access” model.

- **Identify your sensitive information and leverage Microsoft Information Protection and Data Loss Prevention policies to classify and protect it:** As Copilot respects access controls enforced by sensitivity labels and DLP policies, these can prevent oversharing even if users get the permissions wrong. Use DLP and labels to protect sensitive content, and provide user education in the form of policy tips and custom email alerts if a user attempts to perform a prohibited action. Test your access controls before deploying Copilot by using search. If you can find sensitive information in search, Copilot will likely find it too.

- **Assess first- or third-party add-on products that improve visibility and governance features across Microsoft 365:** For large and highly regulated organizations, Microsoft's native controls are often not enough to provide the required visibility and insight across Microsoft 365. This issue will be made worse if Copilot is widely adopted and results in a new era of AI-generated sprawl. To improve your information governance, investigate first-party (from Microsoft) and third-party options that can help to rationalize your Microsoft 365 environment and identify and remediate oversharing risks.

- **Develop mandatory training to help users understand how best to store and share information in Microsoft 365:** Your organization likely has a mandatory security and compliance training program that it repeats multiple times during the year. Develop a module on how best to store, share and collaborate in Microsoft 365 and anchor this to the existing program. Data security and sharing best practices are two sides of the same coin. Reduce risk by ensuring your users know how to share.

- **Promote a "distrust and verify" approach for content generated by Copilot for Microsoft 365 and other AI technologies:** Educate your users to reduce AI-generated mistakes and inaccuracies by helping them to refine their prompts and to provide additional context and information. Be clear that users should not rely on the authenticity and accuracy of AI-generated content and should always validate its sources.

Conclusion

Copilot for Microsoft 365 promises a radical change in terms of how organizations create and discover content, but it can also significantly increase information risks. To get ready for Copilot, and to maximize the potential productivity gains, you must improve your information governance and access controls. Indeed, this should be top of mind whether you plan to purchase Copilot for Microsoft 365 licenses or not.

Evidence

[1.] The Great SaaS Data Exposure, Varonis

[2.] Data, Privacy, and Security for Copilot for Microsoft 365, Microsoft Learn, Microsoft

[3.] Data, Privacy, and Security for Azure OpenAI Service — Azure AI Services, Microsoft Learn, Microsoft

[4] **2023 Gartner Microsoft 365 Survey:** This survey was conducted online from 11 through 31 July 2023 to understand the levels of maturity for Microsoft 365 and how organizations are governing, using and supporting the platform. A total of 150 IT leaders using Microsoft 365 or influencing or making decisions around Microsoft 365 participated — 101 from Gartner's Research Circle (a Gartner-managed panel) and 49 contacted through the dissemination of the survey link via LinkedIn posts and outreach to clients. Respondents who disclosed their locations were in North America (n = 53), EMEA (n = 49), Asia/Pacific (n = 15) and Latin America (n = 6); 27 respondents did not indicate their locations. Disclaimer: Results of this survey do not represent global findings or the market as a whole, but reflect the sentiments of the respondents and companies surveyed.

[5] Poll result from Gartner Client Webinar: GTP Client Webinar: Improve Information Governance to Manage Copilot for Microsoft 365 Risks, 9 November 2023.

Question: What do you think is the biggest risk / issue with deploying Copilot for Microsoft 365? (n = 224)

[6] **Social Media Analysis — Leaders on Generative AI Risks:** Gartner conducts social listening analysis leveraging third party data tools to complement or supplement the other fact bases presented in this document. Due to its qualitative and organic nature, the results should not be used separately from the rest of this research. No conclusions should be drawn from this data alone. Social Media Data in reference is from 1 January 2023 through 31 August 2023 in all geographies (except China) and recognized languages.

Recommended by the Authors

Some documents may not be available as part of your current Gartner subscription.

Quick Answer: How to Make Microsoft 365 Copilot Enterprise-Ready From a Security and Risk Perspective

Solution Path for Implementing Microsoft 365 Governance

Guidance Framework for Managing External Sharing in Microsoft 365

Solution Path for Document Management in Microsoft 365
