



Microsoft Federal Developer Summit Building Al Solutions

Data Security in Preparation for M365 Copilot

Jason Rahn | Security Cloud Solutions Architect | November 13, 2024



AGENDA

- Overview of Microsoft Purview
- Protecting your data with Information Protection
- Interactions with AI
- Information Protection SDK
- Introducing Data with Graph Connector

What does it take to be optimized for AI?

Ready for Microsoft 365 Copilot

Minimum Required: Enable Copilot for EAP users



M365 E3 or E5

- Microsoft 365 Apps for Enterprise (Current Channel)
- · Azure AD account
- · Microsoft OneDrive
- · Microsoft Teams
- New Outlook
- Microsoft Loop

Full requirements available at: <u>Get started with Microsoft 365 Copilot - Microsoft 365</u> admin | Microsoft Learn

Optimized for Microsoft 365 Copilot

Highly recommended: Data compliance, preventing over-sharing and potential internal data leakage





Microsoft Information Protection

M365 E3

- Manual Labelling Policies
- OneDrive and SharePoint Data & Controls
- Microsoft Information Protection

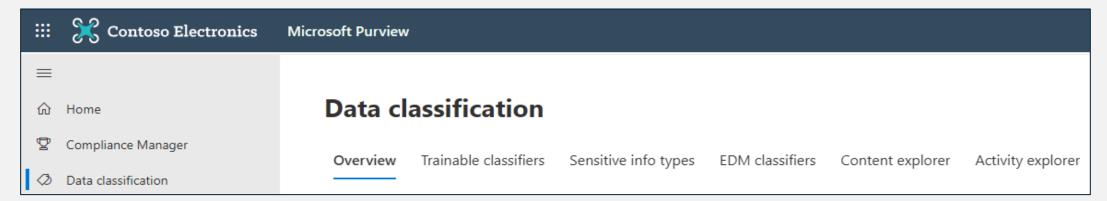
M365 E5

- Automatic Labelling Policies
- Automatic Retention/Deletion Policies
- Access Reviews

Data Classification Section

Data Classification Page is organised into Six Tabs

- Overview
- Trainable classifiers
- Sensitive info types
- EDM classifiers
- Content Explorer
- Activity Explorer



Best-in-class classification technologies



Sensitive Info Types

300+ out of the box info types like SSN, CCN

Clone, edit, or create your own

Supports regex, keywords, and dictionaries

AVAILABLE TODAY



Named

50+ entities covering person name, medical terms, and drug names

Best used in combination with other sensitive info types

AVAILABLE TODAY



Exact Data Match

Provides a lookup to exactly match content with unique customer data

Supports 100m rows and multiple lookup fields

AVAILABLE TODAY



Trainable Classifiers

use trainable classifiers

Create your own classifier based on business data

AVAILABLE IN PUBLIC PREVIEW



Credentials SITs

42 new SITs for digital authentication credential types

Use in auto-labeling and DLP policies to detect sensitive credentials in files

AVAILABLE TODAY



Context-based Classification

Service-side auto-labeling ODSP default site label

- File extension
- Document name contains word
- Document property is
- Document size greater than
- Document created by

ROADMAP ITEM

Templates

- Provide pre-defined policies that use available classifiers
- Cover multiple industry and geographical regulatory requirements

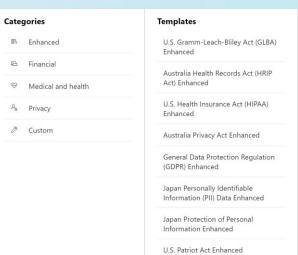
U.S. Personally Identifiable

Enhanced

Information (PII) Data Enhanced

U.S. State Breach Notification Laws

- Easily customizable, can be edited to meet customer needs
- · Get started easily with simulations, rerun as needed to fine tune



U.S. Gramm-Leach-Bliley Act (GLBA) Enhanced

Helps detect the presence of information subject to Gramm-Leach-Bliley Act (GLBA), including information like social security numbers or credit card numbers. This enhanced template extends the original by also detecting people's full names, U.S./U.K. passport number, U.S. driver's license number and U.S. physical addresses. We have also enhanced this template with Trainable Classifier "Business-Finance" and "Business-Tax" which can detect financial information such as Budget proposal, Financial statements and Proposals and Reports and tax information such as Tax planning documents, Tax forms, Tax filing related documents and Tax regulation documents.

Protect this information:

- Business Tax
- · Business Finance
- Credit Card Number
- U.S. Bank Account Number
- U.S. Individual Taxpaver Identification Number
- U.S. Social Security Number (SSN)
- . U.S. / U.K. Passport Number
- . U.S. Driver's License Number
- · All Full Names
- · U.S. Physical Addresses



Why Sensitivity Labels Matter?

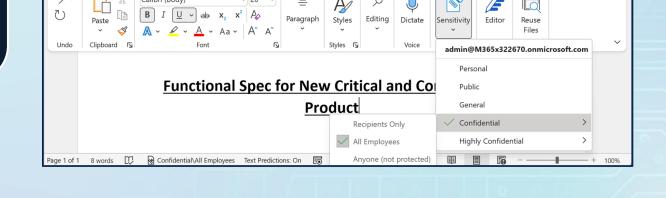
- End-users work with and store data across Office 365 that can be shared with others
- **Some** of this data is sensitive in nature and can present significant risk if stolen, inadvertently shared, or exposed through a breach
- Sensitivity labels can classify and protect this sensitive information regardless of where it exists, even if it leaves your organization
- You can apply sensitivity labels to:
 - Files and emails in Office apps using encryption and content markings
 - Content in third-party apps and services with Microsoft Cloud App Security and the Microsoft Information Protection SDK
 - M365 Groups and SharePoint Sites by with privacy, access, and external sharing controls
 - Data in Power BI, even when it's saved outside the service
 - Assets in Azure, including SQL columns and files in Azure Blob Storage



What is a Sensitivity Label?

A sensitivity label is a stamp that you apply to data to clearly state how sensitive it is and to enforce protection settings. Labels are:

- Customizable and flexible: map your data classification framework to labels to meet business/compliance requirements
- Clear text: label stored in content metadata, allowing third-parties to protect content
- **Persistent:** protection follows email/file regardless of its location



Components of a Sensitivity Label

Label scope

- Files & Emails
- Groups & Sites
- Azure Purview assets (preview)

Protection settings

- Encryption and content markings for files & emails
- Privacy, access control, and external sharing settings for groups & sites

Auto-labeling settings

• Auto-labeling settings (E5/G5) based on identification of trainable classifiers and/or sensitive information types

Labeling data at scale – Guiding principles



Apply labels by default using Label Policy

- When content is created or accessed, set a default label for
 - Files
 - Emails
 - Meetings
 - PBI Reports
 - Containers
 - ...



Apply labels based on context

- By location, for sensitive SPO libraries, site owners can set a default label per library
- For documents at rest
 - Use service-side auto labeling by file size, extension, properties ...
 - Use SetLabel Graph API to label specific files based on your criteria
- For emails in motion
 - Use service-side auto labeling to trigger labels based on predicates like sender, recipient, domain, ...



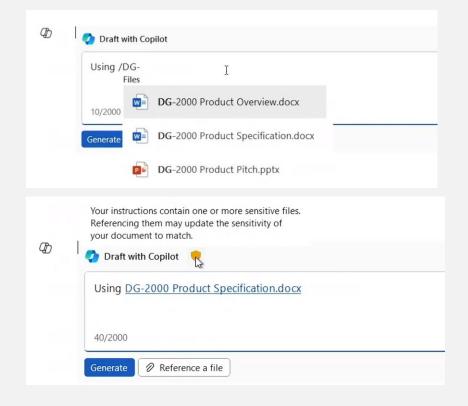
Apply labels automatically based on content

- Client-side autolabeling for content when files are in-use and mails are being composed
- Service side autolabeling policy for files at-rest in SPO, ODB and mails in-transit in Exchange

Content Created using Copilot

When using Copilot to generate new content based on an existing item with a sensitivity label, the sensitivity label from source file is seamlessly inherited, along with its protection settings. Example:

- User employs Copilot in Word, selecting "Draft," then references a file.
- Source content bears the sensitivity label "Confidential\Anyone (unrestricted)" with configured protection settings.
- Result: New content inherits the label "Confidential\Anyone (unrestricted)"





Copilot and Sensitivity Labels (Limitations)

Access Limitations with Encrypted Content:

Copilot functionality may be restricted if encrypted content grants VIEW but not EXTRACT rights.

Label Inheritance Dynamics:

- Inherited sensitivity labels override lower-priority manually applied labels but not higher-priority ones.
- Automatic labeling scenarios follow a similar pattern.

Limitations with Encrypted and Double Key Encryption (DKE) Data:

- Copilot cannot access or return data protected by Double Key Encryption (DKE).
- Encryption configurations may impact Copilot's functionality, particularly with user-defined permissions.

Copilot and Sensitivity Labels (Limitations)

Transition from AIP Add-In:

- The Azure Information Protection (AIP) add-in is retiring in April 2024; Copilot doesn't support it.
- Users are encouraged to migrate to built-in labeling for Office apps for Copilot support.

Teams Sensitivity Labels Limitation:

Copilot doesn't currently recognize sensitivity labels protecting Teams meetings and chats.



Demo Flow Outline

Admin Config

- Classifiers
- Sensitivity Label
- Auto-labeling

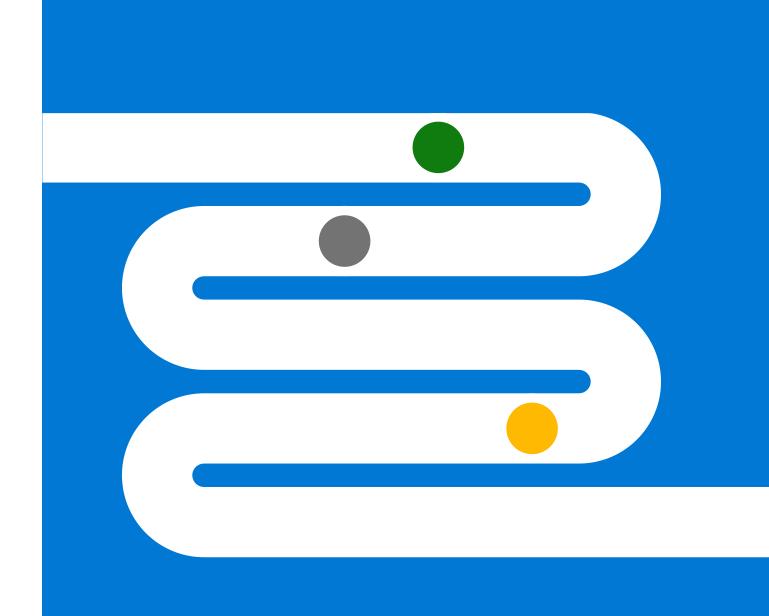
User Experience

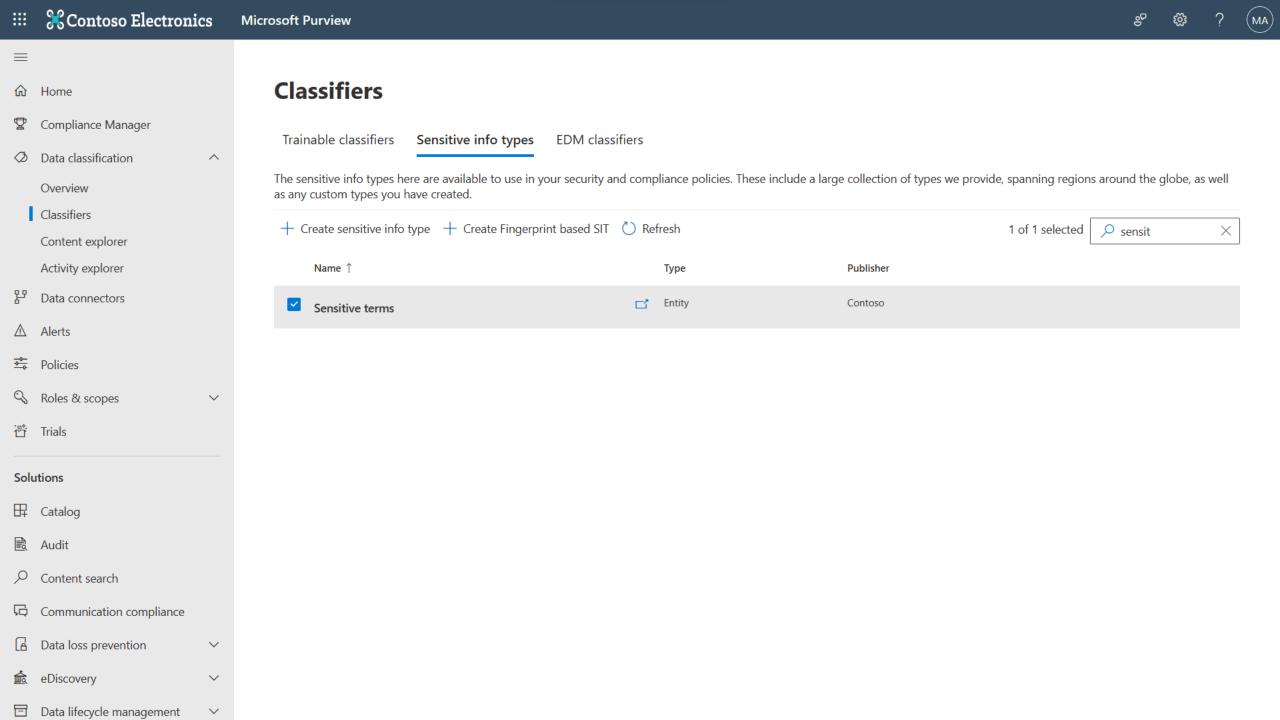
• Alex – Access to the data

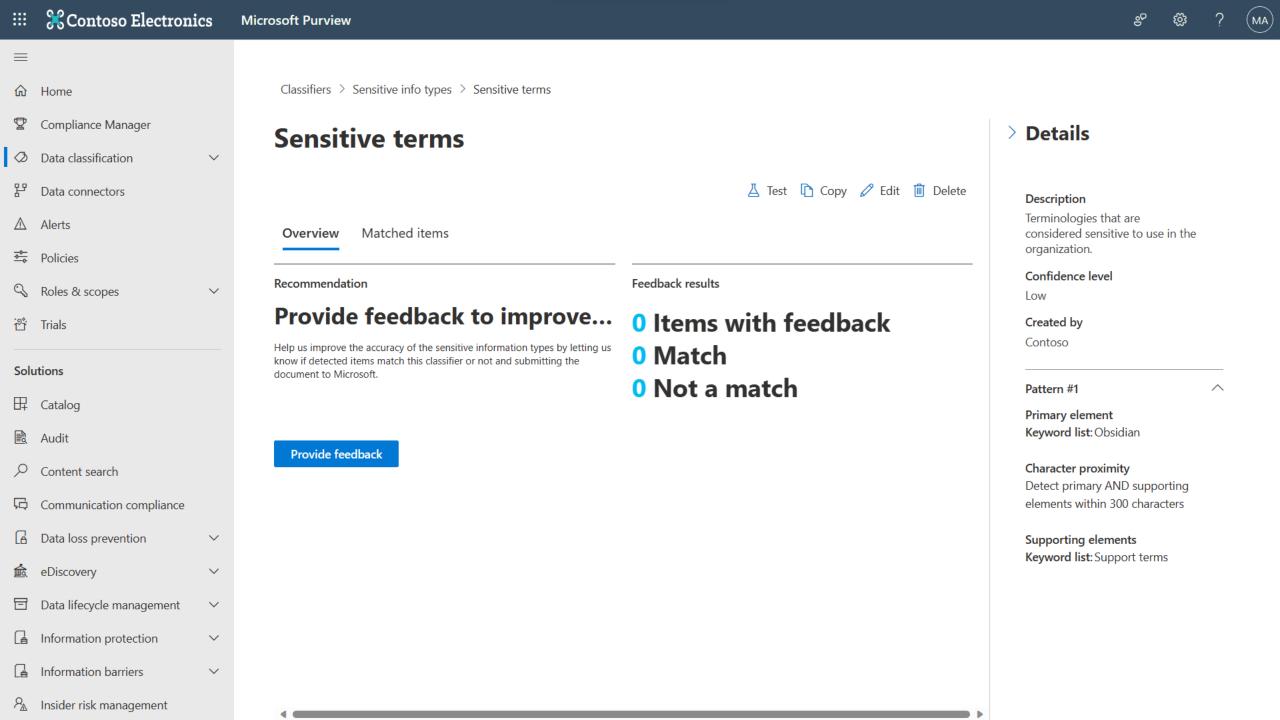
User Experience

• Diego – No access

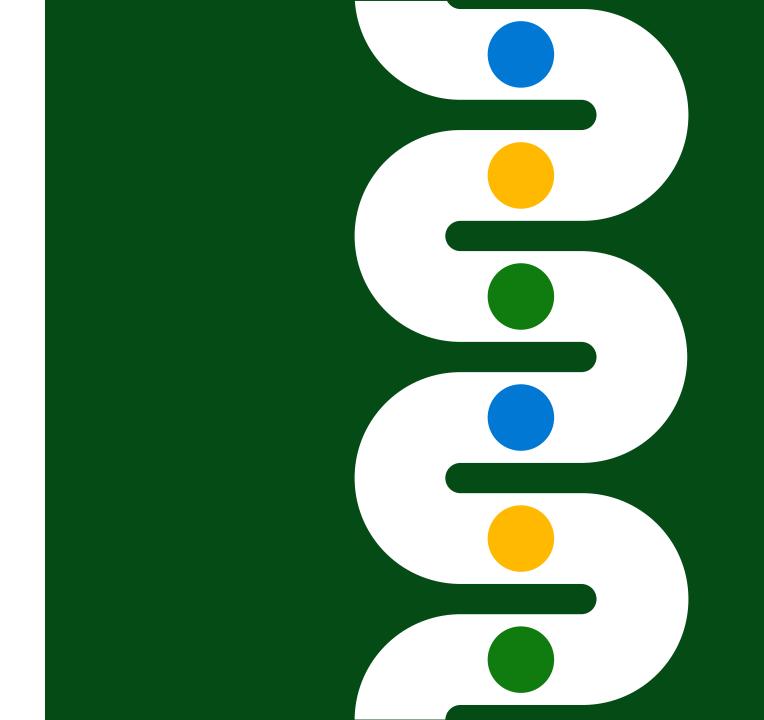
Classifiers

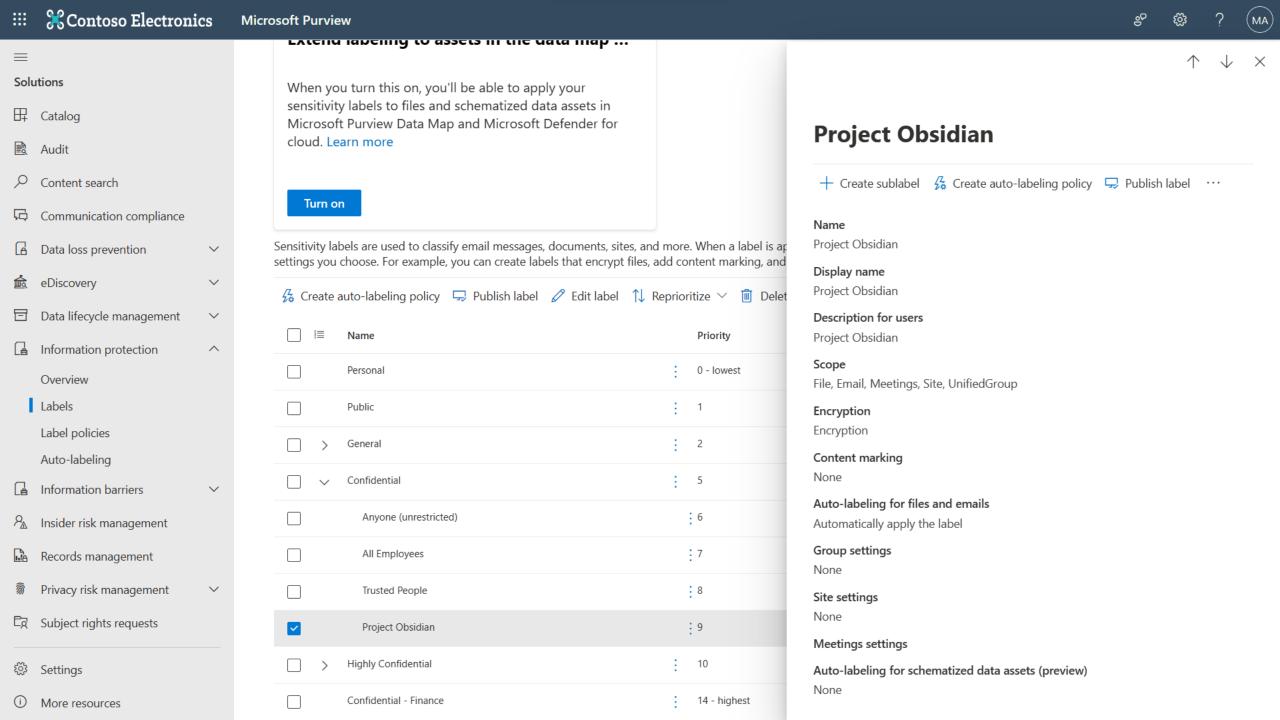






labels, encryption, auto-labeling



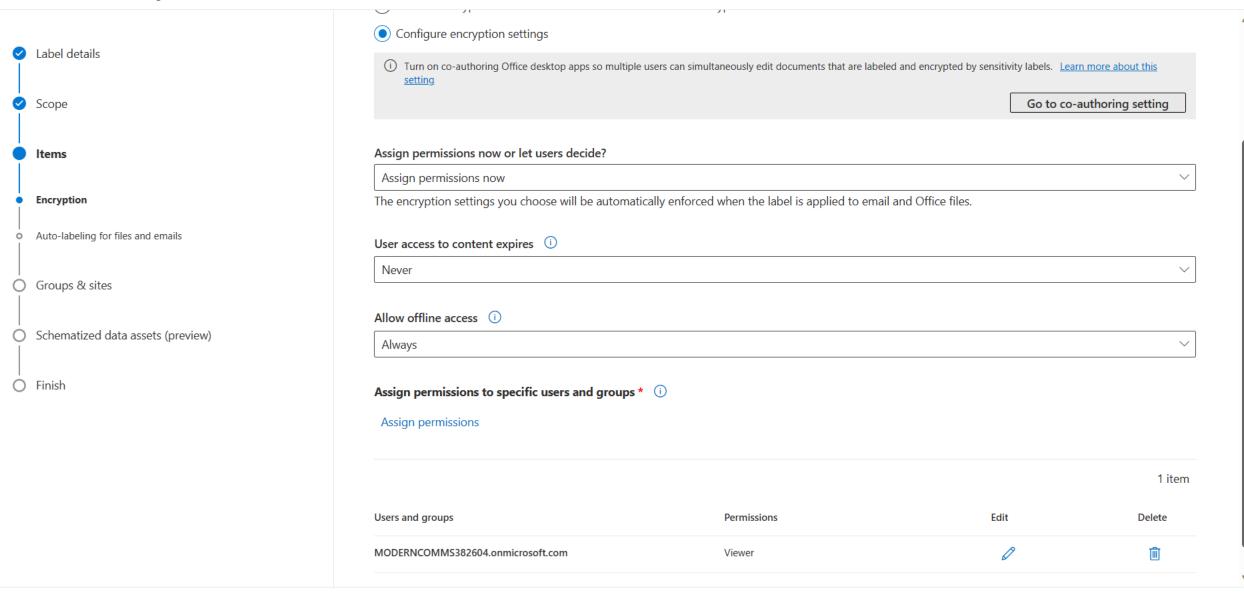








Edit sensitivity label

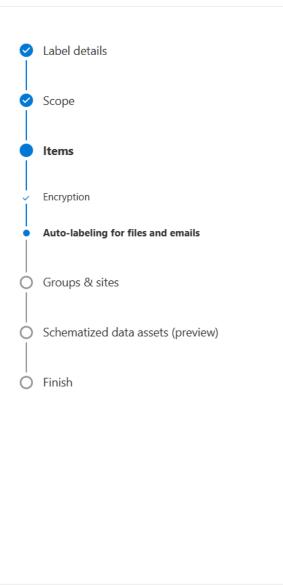


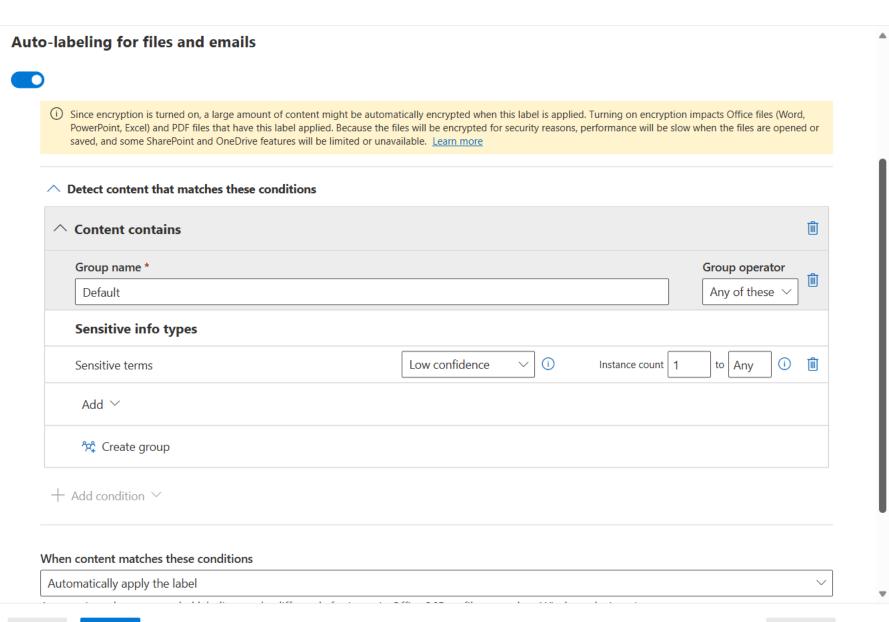




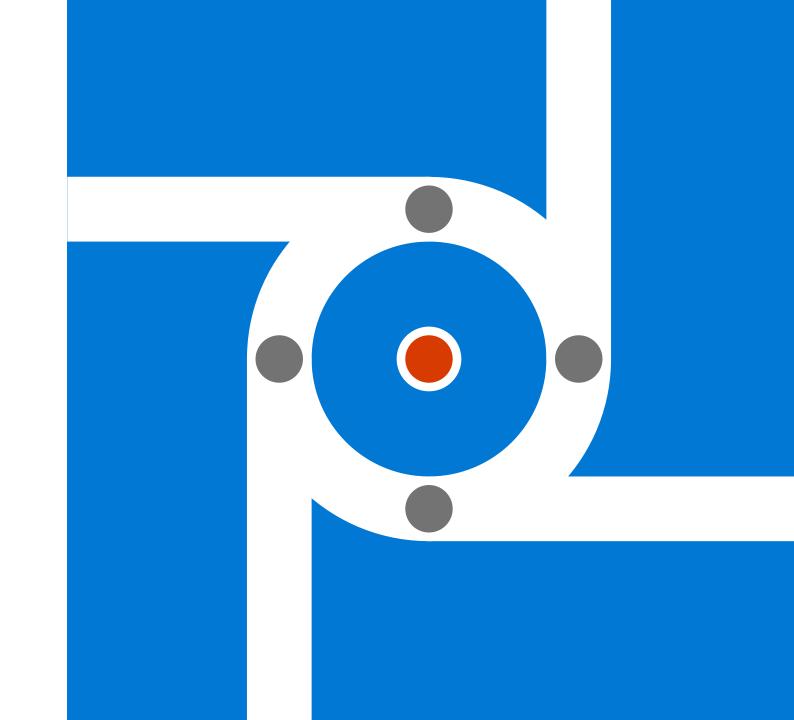


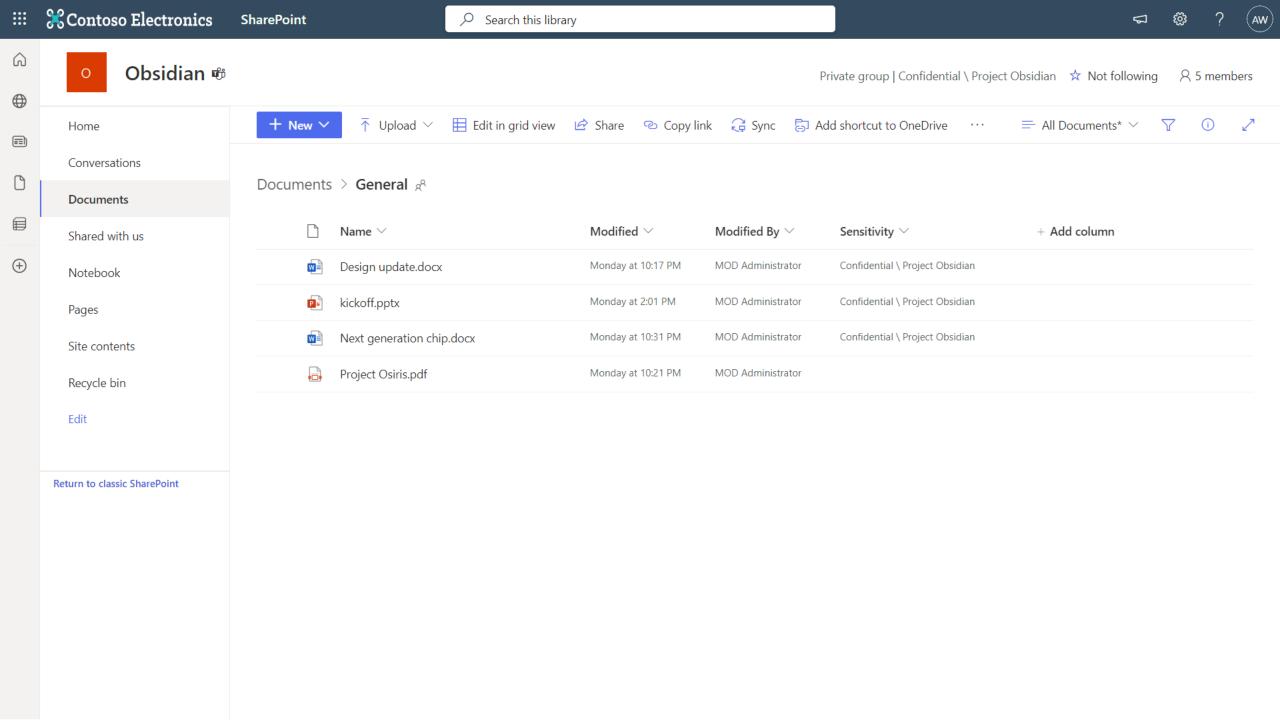
Edit sensitivity label





User experience: Alex











Accessibility Mode





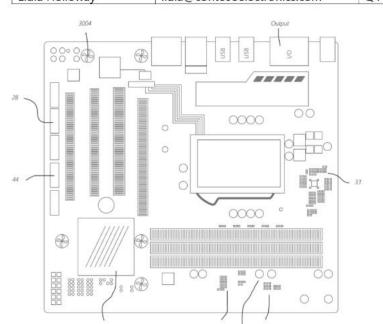


Project Obsidian

Updated Engine Chip Design

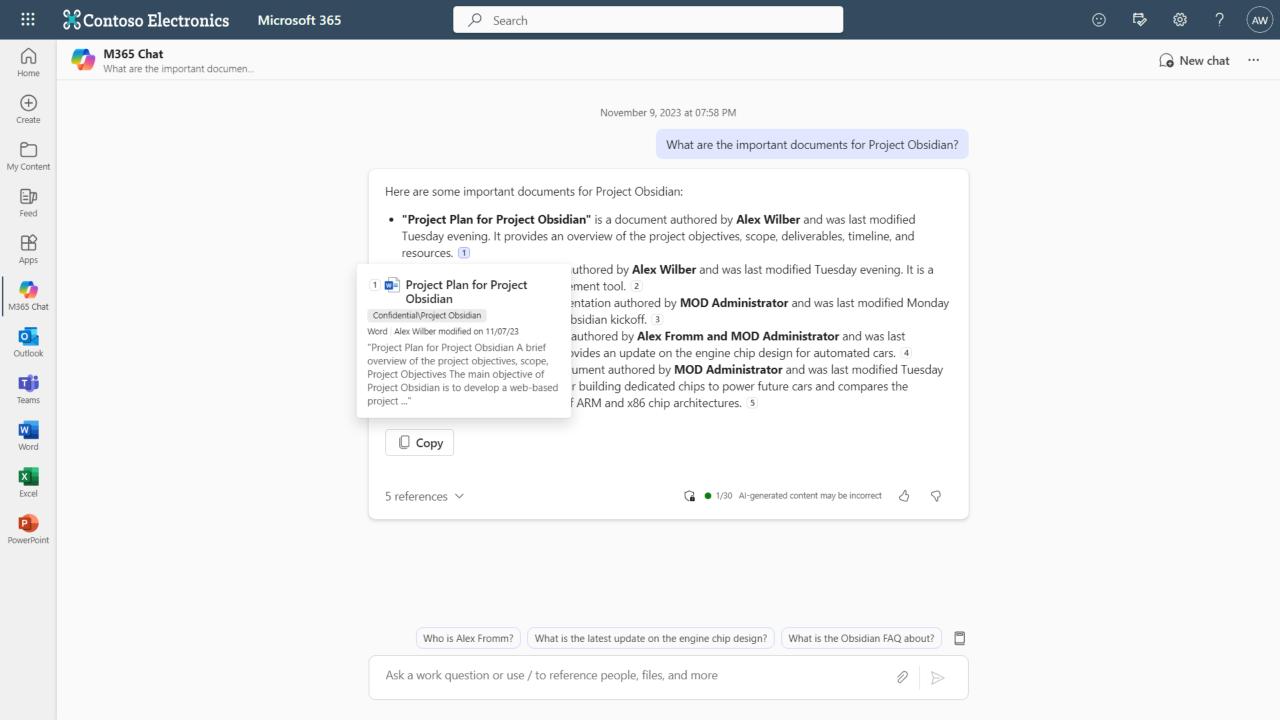
Automated Car Team October 21, 2023

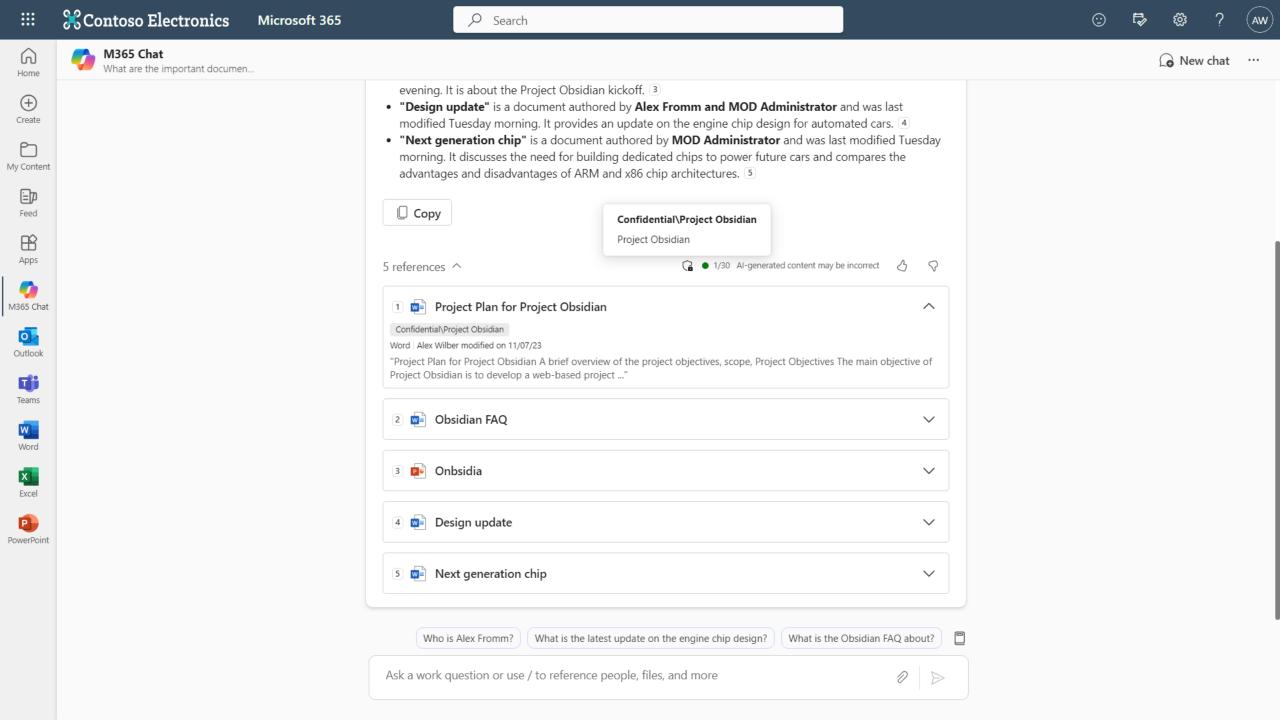
Contacts	Email	Timeline	
Lidia Holloway	lidia@contosoelectronics.com	O4 FY 24	

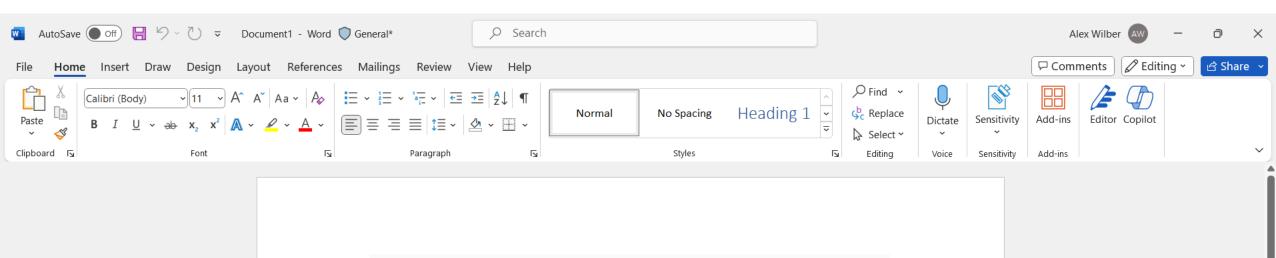


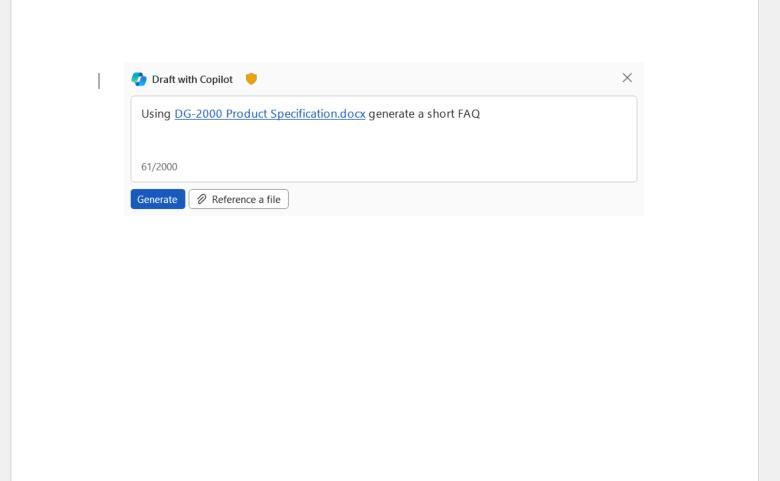
With our new investments in automated cars we need to redesign the Al500 chip to pull more throughput and reduce overheating.

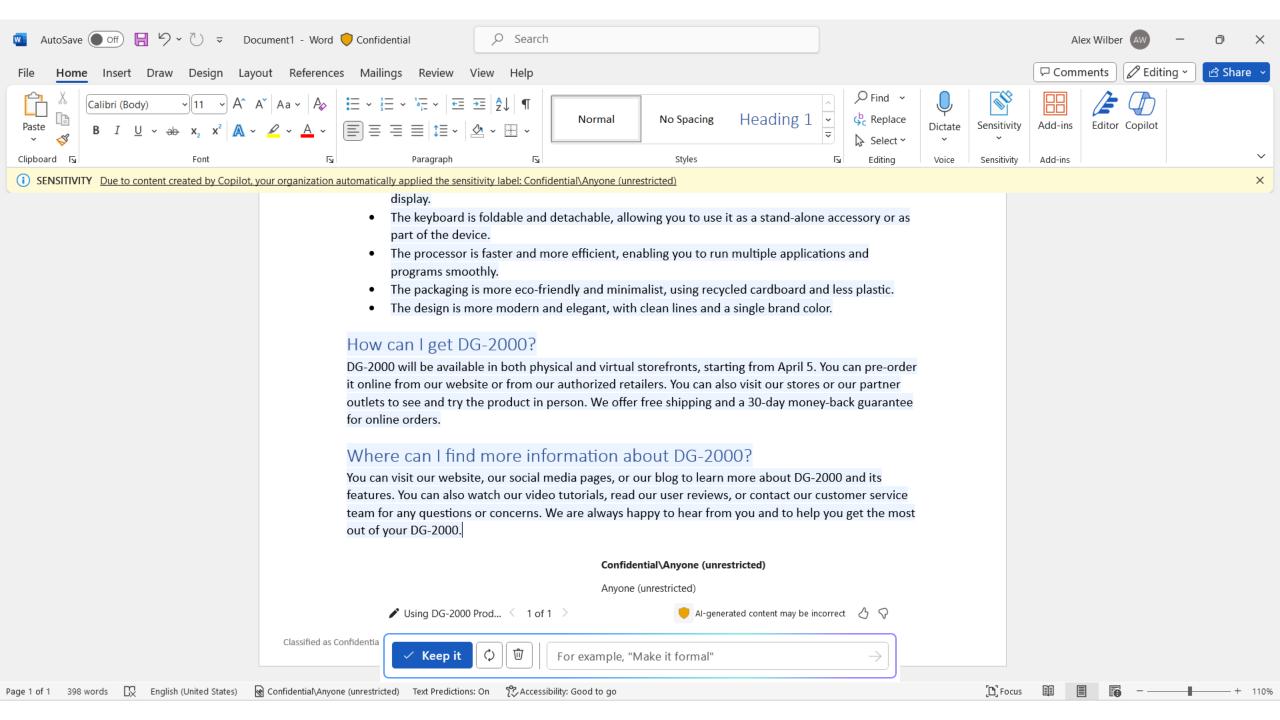
We have attached more fans to the base chip but need more focus on the transistors across the entire board.

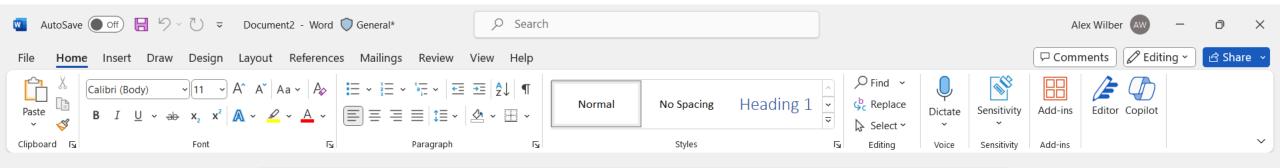


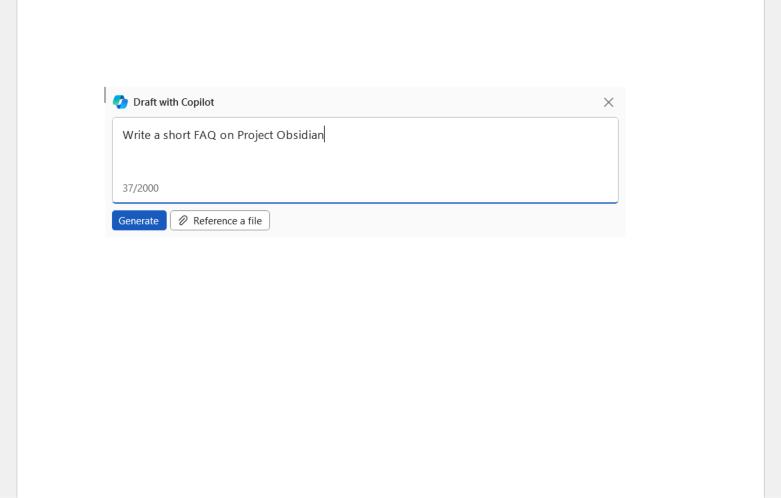


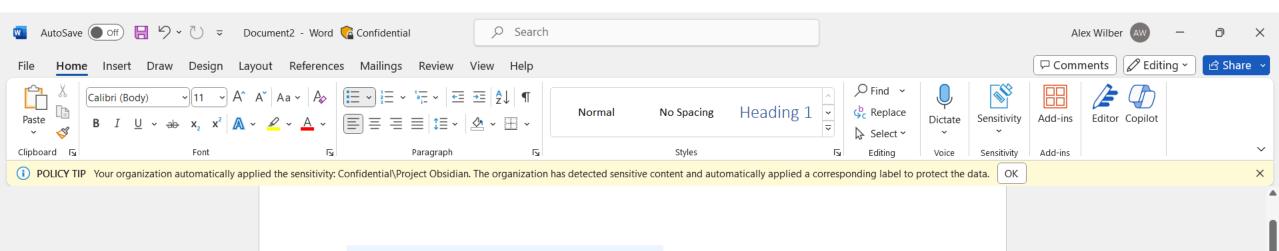










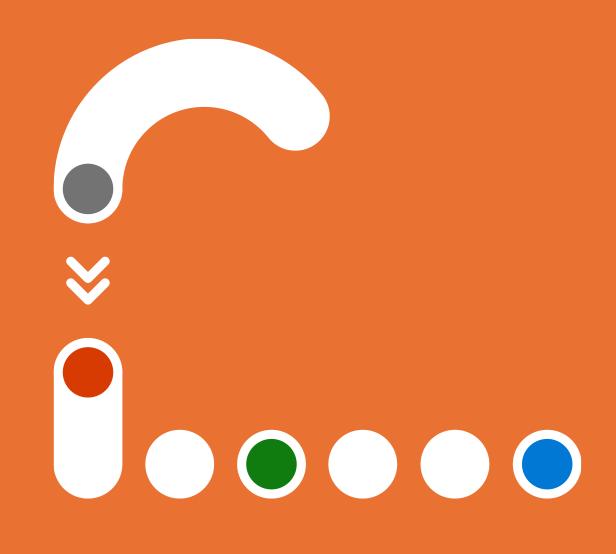


FAQ on Project Obsidian

- What is Project Obsidian?
- Project Obsidian is a research initiative that aims to develop a novel framework for secure and scalable distributed computing. The project is led by a team of experts from academia and <u>industry</u>, and is funded by a grant from the National Science Foundation.
- What are the main goals of Project Obsidian?
- The main goals of Project Obsidian are to:
- Design and implement a new protocol for consensus and fault tolerance in large-scale networks
- Develop a new programming language and runtime for writing and executing distributed applications
- Provide a platform for testing and benchmarking the performance and security of distributed systems
- Explore the applications and implications of distributed computing in various domains, such as cloud computing, blockchain, IoT, and machine learning
- How can I get involved in Project Obsidian?
- Project Obsidian is an open and collaborative project that welcomes contributions from anyone
 who is interested in distributed computing. You can get involved in Project Obsidian by:
- Inining the project mailing list and Slack channel to stay undated and communicate with the



User experience: Diego

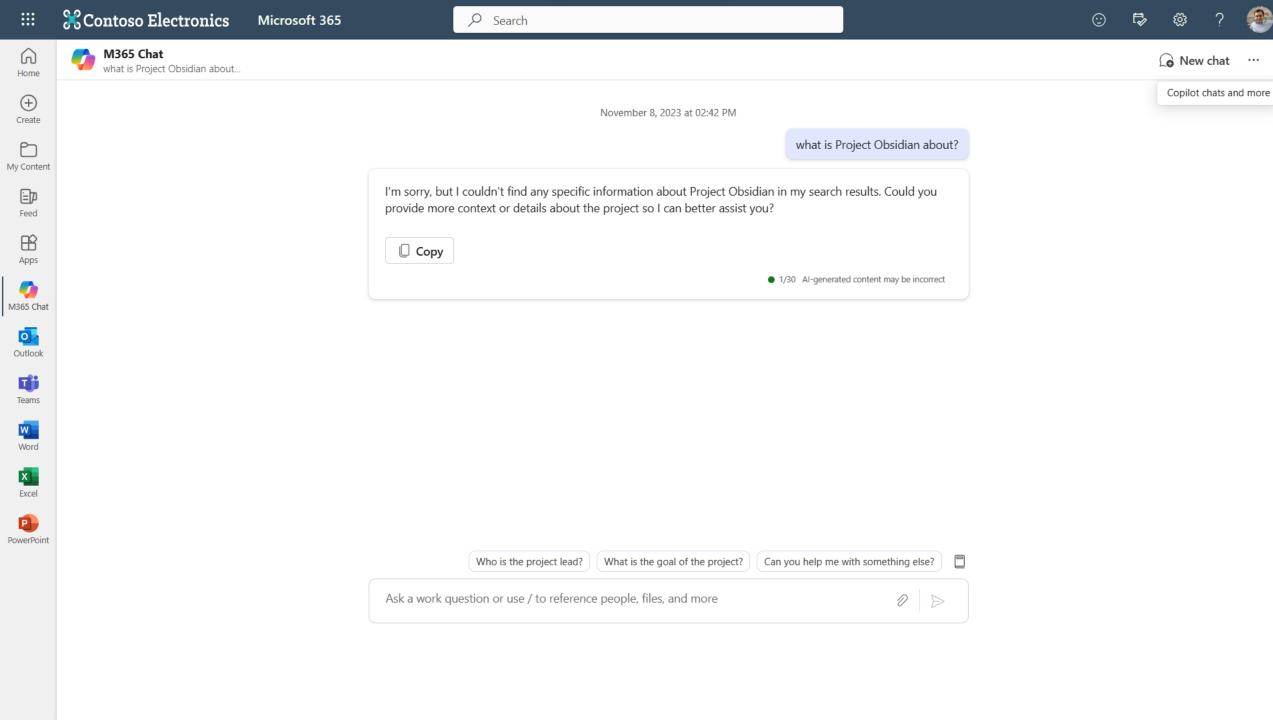


Sorry, you don't have access.

I'd like access, please.

Request Access

Diego tries to access the Project Obsidian site



Microsoft Information Protection SDK

File API

Policy API

Protection API

https://aka.ms/mipsdkbinaries

https://aka.ms/MIPSDKDocs



File API

List Sensitivity Labels

Apply/Change/Delete label on file

Downgrade label and record justification

Republish File

Supported File Types:

- •Adobe Portable Document Format: .pdf
- •Microsoft Power BI: .pbit, .pbix
- •Microsoft Project: .mpp, .mpt
- •Microsoft Publisher: .pub
- •Microsoft XPS: .xps .oxps
- •Images: .jpg, .jpe, .jpeg, .jif, .jfif, .jfi, .png, .tif, .tiff
- •Autodesk Design Review 2013: .dwfx
- •Adobe Photoshop: .psd
- •Digital Negative: .dng
- Microsoft Office

Policy API

Lists available labels for current user

Lists actions that should be applied

The application must enforce policy and take any actions on the file



Protection API

List Protection Templates Encrypt/decrypt files

Microsoft Graph Connectors



Connect to MSGraph

Microsoft Graph

Proprietary External Connection Microsoft 365

M365 Data Source/Vertical



Steps to create a connector

App ID

App Registration created in Entra ID

Permissions needed:

- ExternalConnecti on.ReadWrite.Ow nedBy
- ExternalItem.Rea dWrite.OwnedBy

Create the connection

Id – used toidentifyconnector inGraph API calls

Name – self explanatory

Description – self explanatory

Configuration

(optional) – List of Entra App Ids that can manage the connection

Create the schema

baseType must be Microsoft.graph.e xternalItem

Create a collection of properties with the following data:

- Name
- •Type data type
- •isSearchable boolean
- •isRetrievable boolean

Send data

All items are externalData items and must match the schema

Publish Vertical

This allows various M365 Experiences to use this data



Graph API calls – Create Connection

Url	Method	Body
http://graph.Microsoft.com/v1.0/external/connections	POST	{ "id": "contosohr", "name": "Contoso HR", "description": "Connection to index Contoso HR system" }

```
Header:
{
    "Authorization": "Bearer {Token}",
    "Content-Type": "application/json"
}
```



Graph API calls - Create Schema

Url	Method	Body
http://graph.Microsoft.com/v1.0/external /connections/contosohr/schema	PATCH	{ "baseType": "microsoft.graph.externalItem", "properties": [{ "name": "title", "type": "String", "isSearchable": "true", "labels": ["title"] }, { "name": "url", "type": "String", "isQueryable": "true", "isRetrievable": "true", "isSearchable": "false" }]

External Item

ID ACL Content properties activities

ID

• developer provided String

ACL

- Type
- Value
- accessType

Content

Plain text representation of content

Properties

 Must conform to schema that was provided

Activities

- View
- Modified
- Created
- commented



Graph API calls - Create externalltem (Send Data)

Url	Method	Body
http://graph.Microsoft.com/v1.0//external/connections/contosohr/items/appFile12345	PUT	<pre>{ "acl": { "type": "user", "value": "e811976d-83df-4cbd-8b9b-5215b18aa874", "accessType": "grant" }, "properties": { "title": "Error in the payment gateway", "url": "https://myapp.contoso.com/rawfiles/subdirectory/appfile12345" }, "content": { "value": "Text representing the data within the file", "type": "text" } }</pre>

Graph API calls – Add Activities

Url	Method	Body
http://graph.Microsoft.com/v1.0//external/connections/contosohr/items/appFile12345/addActivities	POST	{ "activities": [{ "@odata.type": "#microsoft.graph.externalConnectors.externalActivity", "type": "created", "startDateTime": "2021-04-06T18:04:31.033Z", "performedBy": { "type": "user", "id": "1f0c997e-99f7-43f1-8cca-086f8d42be8d" } } } }





Considerations for Graph Connectors

Application sets the ACL

Unique IDs are controlled by the application

Your application must make all updates to the data

Add User activities to your items

Set your vertical to show results inline



Apply Semantic Labels

title	The title of the item that you want shown in search and other experiences.		
url	The target URL of the item in the data source.		
createdBy	The name of the person who created the item in the data source.		
lastModifiedBy	The name of the person who most recently edited the item in the data source.		
authors	The names of all the people who participated/collaborated on the item in the data source.		
createdDateTime	The date and time that the item was created in the data source.		
lastModifiedDateTime	The date and time that the item was last modified in the data source.		
fileName	In case of a file, the name of the file in the data source.		
fileExtension	In case of a file, the extension of the file in the data source.		
iconUrl	The URL of an icon.		
containerName	The name of the container.		
containerUrl	The URL of the container.		

References

Learn about sensitive information types | Microsoft Learn
Learn about sensitivity labels | Microsoft Learn
Encryption in Microsoft 365 | Microsoft Learn
Microsoft 365 Copilot overview | Microsoft Learn

Overview - Microsoft Information Protection SDK. | Microsoft Learn

- Quickstart Initialization for Microsoft Information Protection (MIP) SDK C# clients | Microsoft Learn
- Quickstart Client application initialization Protection SDK (C#) | Microsoft Learn

Use the Microsoft Graph connectors API - Microsoft Graph v1.0 | Microsoft Learn

- externalConnection resource type Microsoft Graph v1.0 | Microsoft Learn
- schema resource type Microsoft Graph v1.0 | Microsoft Learn
- externalltem resource type Microsoft Graph v1.0 | Microsoft Learn

Microsoft Graph | Microsoft Learn



Contact Us

If you have any questions, please feel free to e-mail us here:

FedCSUDataSecurity@microsoft.com

THANK YOU!

Jason Rahn |

Jason.Rahn@microsoft.com |





Feedback

Do you want us to follow up after the event? Do you have feedback?



https://aka.ms/summit/feedback



