AI and Security FAQ

Version: Therefore™ 2026 (35.0.0)

Date: November 05, 2026

© 2026 Therefore Corporation GmbH

www.therefore.net

Introduction

This document was formerly known as the Therefore™ Data Extraction and Classification FAQ.

Therefore™ offers cloud-based, AI powered solutions for classifying and extracting data from documents. These solutions are configured in the Therefore™ Solution Designer, under the Smart Capture node.

Furthermore, Therefore™ offers an integration with Microsoft Copilot, allowing users to transmit data to Copilot for further processing.

Due to these features' cloud-based nature, we at Therefore™ understand that our partners and customers may have many questions about their operation and security. The purpose of this document is to address some of these and provide more information on these aspects.

Document Extraction

This is a service run by Therefore Corporation to assist customers in extracting data from invoices they upload to their Therefore™ system. To start this process, invoices must be saved to a designated workflow that contains the tasks required to process it. During processing, the documents are uploaded to a recognition server. This server reads the data on the invoice and presents it to a Therefore™ user, who must manually check the results and approve each invoice. The manual verification step may be skipped if the administrator configures a confidence level, whereby the Smart Capture engine makes a decision as to whether it read the data on the document correctly or not. If it scores the document below the confidence level, it will send the document to a user for manual approval.

Is the usage of Smart Capture for extraction optional?
Yes, Smart Capture is a completely optional feature. To use Smart Capture, the Therefore™ system must support workflow (Essentials, Business, Enterprise, or Online Edition) AND Smart Capture pages must be purchased before the feature can be activated. More information on licensing can be found in separate documentation.

Are all documents in the Therefore™ system impacted by Smart Capture?
No. Only those documents that are sent through the designated workflow for uploading to Smart Capture, whether through an automated or manual workflow trigger, are impacted.

Therefore™ Online customers: Is the processing done within the Therefore™ Online environment?
New invoices are first saved in Therefore™ Online. Once uploaded to Smart Capture, the invoice leaves the Therefore™ Online environment. The recognition server is hosted in a data center outside of the Azure-based Therefore™ Online center. After processing, the extracted invoice data is sent back to the Therefore™ Online environment. See the next section ‘Data center’ for more details about the recognition server’s data center.

Therefore™ On-Premises customers: How does Smart Capture work with on premises systems?

To use the features of Smart Capture with Therefore™ On-Premises, the system must be running at least Therefore™ 2021 and must be connected to the internet. When a saved invoice is uploaded to Smart Capture, the invoice leaves the Therefore™ Server. The recognition server is hosted in an external datacenter (note: NOT the Therefore™ Online environment). After processing, the extracted invoice data is sent back to the Therefore™ Server. See the next section 'Datacenter' for more details about the recognition server's datacenter.

Where does the processing take place?

Therefore™ Smart Capture leverages an artificial intelligence service run on a recognition server hosted on Amazon Web Services, a leading cloud infrastructure provider. The AI service is operated by a trusted and vetted technology partner under contract with Therefore Corporation.

Where are the recognition servers hosted?

There are two recognition servers.

The recognition server for customers based in the United States is hosted in an AWS datacenter located within the United States. The recognition server for customers in all other regions is hosted in an AWS datacenter located in the Republic of Ireland, within the European Union.

Is the service GDPR compliant?

Therefore Corporation is fully committed to ensuring compliance with the GDPR. Therefore™ software has passed a rigorous data protection audit by the consulting firm Ebner Stolz and its solutions have been verified and certified to enable an effective management of data in accordance with GDPR regulations. By running the service in an EU-based datacenter, the AI service provider is equally bound by and subject to the GDPR.

How long are documents retained on the recognition server?

The standard retention period for documents sent to the recognition server is 365 days.

However, the retention period can be changed to a custom number of days within the Smart Capture configuration settings in Therefore™.

Data is retained to further develop and refine the data extraction capabilities of the Therefore™ Smart Capture service. The data is encrypted both in transit and at rest. Each customer's data is used for learning on their local tenant only; data is never shared with other tenants nor used for any type of global learning.

Data is deleted after the retention period. This means that any learning from data processed before the end of the retention period is also lost. The following examples illustrate this point. In this scenario, we are assuming that the retention period is the default value (365 days or 12 months).

- You process an invoice from supplier 'Moyaware' on June 1st 2023. You manually correct this invoice, allowing the system to learn from the correction. On June 1st 2024, the invoice and any related learning is deleted from the system. If you process another invoice of the same type from 'Moyaware' any time after June 1st 2024, you may have to correct it again.
- You process invoices from supplier 'Moyaware' on June 1st 2023 and February 1st 2024. On June 1st 2024, the first invoice (from June 2023) and any related learning is deleted from the system. However, the system still retains the learning from the second invoice (from February 2024), since this invoice will be deleted on February 1st 2025.

To summarize: Any learning from an invoice type (same issuer and format) is deleted alongside the invoice after the end of the retention period. However, if other invoices of that same type are continually processed within the retention period after the first, the learning retention will be extended. Specifically, the learning is retained for the length of the retention period after the last invoice of that type is processed.

How is data encrypted?

Data sent to and from the recognition server is always encrypted with TLS v1.2 using https. Data at rest on the recognition server is encrypted using AES256. This means that

the data is protected from hacking or unauthorized access by industry-leading technologies both while in transit and while stored on the server.

Is the Smart Capture data center ISO 27001 compliant?

Yes. The AWS data center in Ireland where invoices are processed through Smart Capture is certified as ISO 27001 compliant. Furthermore, the AI-based recognition service itself operated under contract with Therefore Corporation has independently achieved ISO 27001 compliance. This means both the hosting infrastructure and the AI service processing the invoices are ISO 27001 certified.

In addition to ISO 27001, Smart Capture's AI processing platform holds:

- ISO 42001 - the international standard for AI Management Systems, covering responsible AI governance.
- SOC 2 Type II — independently audited across all five Trust Service Criteria: Security, Availability, Processing Integrity, Confidentiality, and Privacy.
- GDPR Compliant — fully compliant with the EU General Data Protection Regulation.
- HIPAA Compliant - meets healthcare-grade data protection requirements.

Does Smart Capture use customer invoices to improve its AI?

Yes - but only to improve accuracy for your organization and not for the benefit of other customers. Smart Capture operates with a two-tier learning model:

- Global base models - pre-trained on a broad set of invoice formats before deployment.
- Your private learning layer - Smart Capture continuously learns from your organization's invoices and user corrections, improving extraction accuracy over time for the organization's specific suppliers and invoice formats.

Your invoice data is not used to train AI models for other customers. Your learning layer is private and isolated to your organization.

How does the AI learning process work in practice?

When your team reviews and corrects extracted data in Smart Capture, this human

feedback improves the model's accuracy for future invoices from the same supplier. The learning is automatic and continuous. Over time, invoices require progressively less manual correction. Smart Capture also retains invoices in processing queues (with extracted data and original files) within the configured retention period to support audit, reprocessing and AI improvement.

Is the invoice data shared?

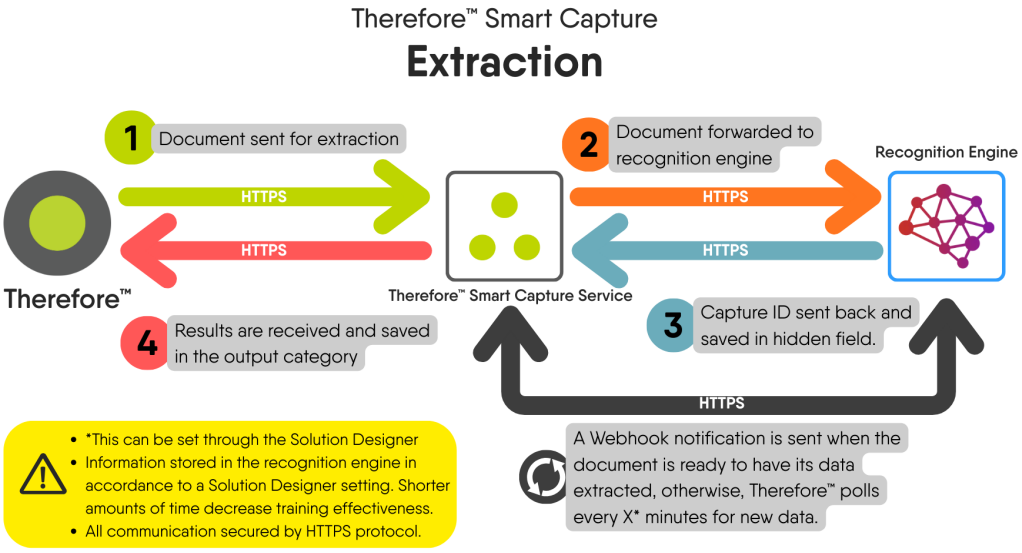
Invoice data is not shared with other organizations. Your data is isolated. Customer data is logically separated at rest, with strict security filters on all database queries. Smart Capture does not sell customer data, does not expose invoice data externally, and enforces strict internal access policies.

The table below summarizes the key security and compliance measures that protect invoice data processed through Therefore™ Smart Capture

Area	Measure / Standard
Smart Capture Usage	Optional feature, only documents routed through the designated Smart Capture workflow are affected.
Data Center Location	AWS data center, Republic of Ireland, European Union: ISO 27001 certified. For US systems, a US-based AWS data center is used, also ISO 27001 certified.
GDPR and ISO Compliance	Therefore™ has been independently audited for GDPR compliance and is ISO 27001 certified.
Data Architecture	Smart Capture: temporary processing layer; Therefore™: permanent record system.

Area	Measure / Standard
Encryption in Transit	TLS v1.2 via HTTPS: all data encrypted end-to-end between Therefore™ and the recognition server.
Encryption at Rest	AES-256: all data stored on the recognition server is encrypted.
ISO 27001	Certified: the AWS data center and the AI recognition service independently hold ISO 27001.
SOC 2 Type II	Independently audited across all five Trust Service Criteria.
ISO 42001	AI Management System Standard: responsible AI governance.
AI Learning	Private learning layer per organization: your data is never used to train other customers' models.
Access Control	Role-based access, fine-grained permissions, audit logging; no external data exposure.
Retention Period	Configurable in Smart Capture (default: 365 days); permanent storage in Therefore™

Diagram of Therefore™ Smart Capture Extraction



Document Classification

This is a service run by Therefore Corporation to assist customers in classifying batches of documents into their respective categories according to type. This feature uses Microsoft Azure Document Intelligence to classify documents and assign them to their respective categories.

From a technical perspective, the feature works as such:

- Documents saved to a classification category are automatically classified and moved to the final category. This happens within a workflow.
- The classifier engine can be trained based on existing documents in the final categories (minimum 5 documents per category, although more documents will improve classifier accuracy).
- The workflow then moves documents to the destination category.
    - Manual workflow step in case the classifier confidence is low.
- Documents can then be processed normally/in any other way, including being sent for data extraction with Therefore™ Smart Capture.

In the case of multi-page documents, administrators can configure the number of pages of a document that should be used for classification. For example, setting this to "1" means the classifier will only read the information on the first page to classify the document. "2" means it will read the first two pages.

Important note: this feature is about document classification, not extraction. It may be used in conjunction with Therefore™ Smart Capture extraction (in a separate workflow) to extract data from documents that were classified by Azure Document Intelligence.

Is the usage of Document Classification optional?

Yes, Document Classification is a completely optional feature. To use it, the Therefore™ system must be running in Therefore™ Online or Therefore™ Subscription Edition. A connector license for Azure Document Intelligence is required. Document classification pages must be purchased before the feature can be activated. Packs of pages must be

purchased in advance, similarly to Therefore™ Smart Capture. Pack sizes are 100 / 1.000 / 10.000. Annual subscription models also have an option for a pack of 100.000 pages.

Are all documents in the Therefore™ system impacted by Document Classification?
No. Only those documents that are sent through the designated workflow for uploading to Azure Document Intelligence, whether through an automated or manual workflow trigger, are impacted.

Therefore™ Online customers: In which environment is the processing done?
Documents sent for classification with Azure Document Intelligence stay in the same regional datacenter as the customer's tenant. For example, the documents a customer from the United States would be processed through Azure Document Intelligence in the US-based Azure datacenter. The documents of a customer from Germany would be processed in the EU-based Azure datacenter.

Therefore™ On-Premises customers: How does Document Classification work with on premises systems?
To use the Document Classification features of Azure Document Intelligence with Therefore™ On-Premises, the system must be running at least Therefore™ 2025 and must be connected to the internet. When documents are uploaded to Azure Document Intelligence, they leave the Therefore™ Server. The processing takes place in the local Therefore™ Online environment, as described above. After processing, the classification data is sent back to the Therefore™ Server.

Where are documents uploaded to when using Azure Document Intelligence? Where are they retained and for how long?
There are two types of storage to consider: training documents and classification documents.

Training documents are the reference documents that the classification engine learns from. These are initially uploaded when training a new classifier and are also added whenever a manual correction is done by a user. These training documents are stored indefinitely in the Therefore™ Online storage of the corresponding regional Azure

datacenter to retain the classifier's training results and improve its performance. Training data is not grouped between customers but rather separated according to the Customer ID.

Classification documents are documents that are sent for classification through a workflow to Azure Document Intelligence. The processing happens in the corresponding regional Azure datacenter. According to Microsoft documentation, the input data and results are deleted within 24 hours and not used for any other purpose.

How is data encrypted at rest and in transit when using Azure Document Intelligence?

During transit, all data uses https communication to ensure the requests are encrypted. Once the request arrives at the Therefore™ Online datacenter, all transfers are done securely inside the Azure data center infrastructure.

At rest, training data is stored within Therefore™ Online using Azure Secure Storage Encryption (SSE). Further details from Microsoft can be found here:
Azure Storage encryption for data at rest | Microsoft Learn

At rest, classification data also uses Azure Secure Storage Encryption. This is managed by Microsoft in Azure Document Intelligence according to the details outlined here:
Data, privacy, and security for Document Intelligence - Azure AI services | Microsoft Learn

Do customers have any configuration settings that allow them to define how Azure Document Intelligence processes documents?

No. Customers can only configure how many pages of a document are sent for training or classification.

Diagram of Therefore™ Smart Capture Classification - Training

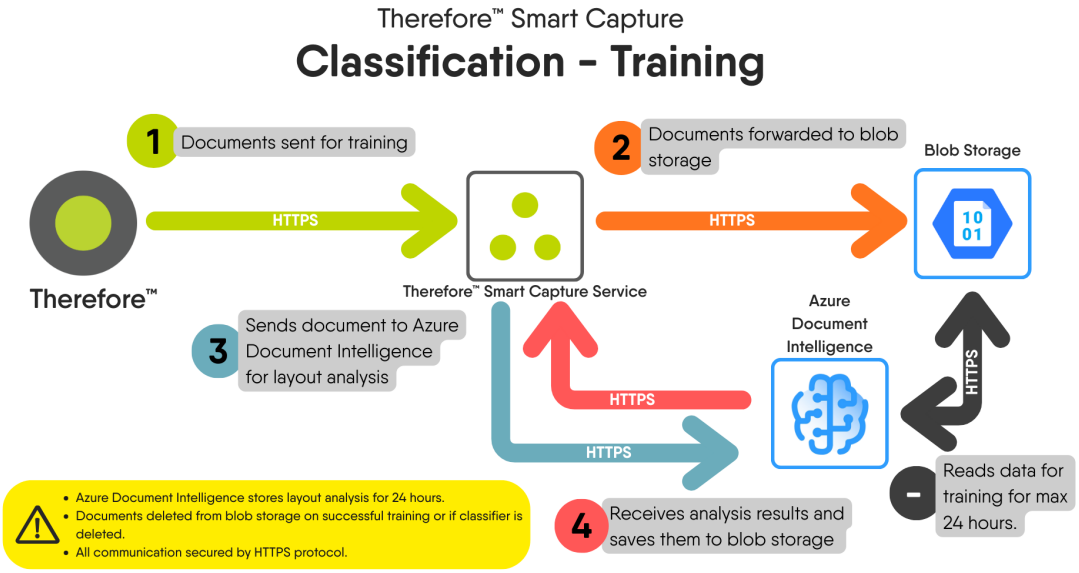
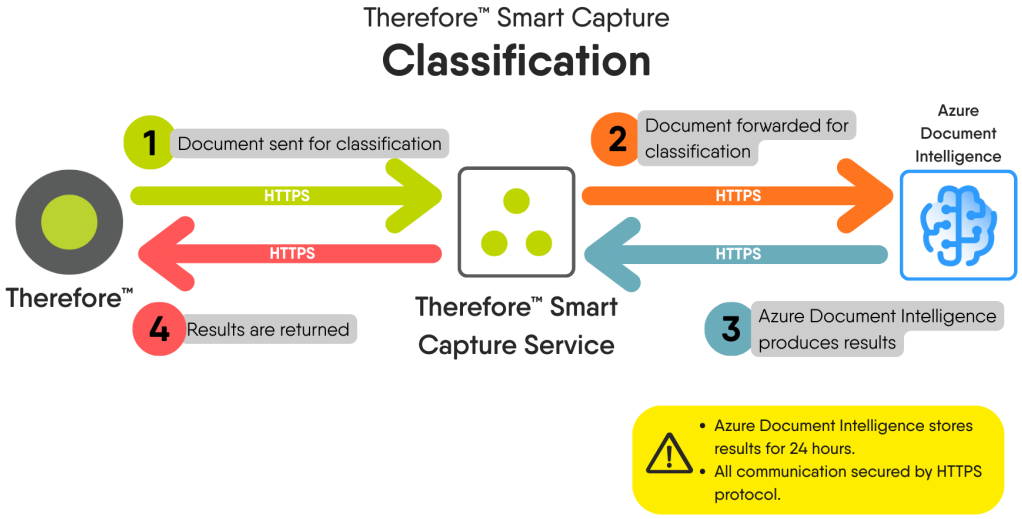


Diagram of Therefore™ Smart Capture Classification



Document Insights (Copilot)

Therefore™ offers an integration with Microsoft Copilot, allowing you to leverage AI to gain useful insights about the information you store in Therefore™. From a technical perspective, this feature works as such:

- Proper planning within the customer's Azure environment is required. As a prerequisite, Copilot must be configured and set up in the customer's own Azure environment. This requires several preparation steps, such as registering a new application and assigning the proper API permissions. Details may be found in the Therefore™ Online Help or in Therefore™ eLearning courses.
- A configuration dialog in the Therefore™ Solution Designer allows the administrator to configure a Microsoft Graph connection.
- The administrator will also choose which Therefore™ categories should be uploaded to Copilot via the Microsoft Graph connection. The Therefore™ Connector uploads the information from the documents in the selected categories to Microsoft Graph.

Is this feature optional?

Yes. The Copilot integration is a completely optional feature. To use it, the Therefore™ system needs a Copilot Connector license (purchased from Therefore), as well as Copilot licenses for users (purchased from Microsoft).

Are all documents in the Therefore™ system affected by this integration?

No. Only those documents in categories which the administrator chooses to add to the Graph connection are affected.

In which environment is Copilot running?

Copilot runs in the customer's own Azure environment. Therefore™ does not provide this environment. As explained in the introduction above, the customer must provide their own Azure environment with Copilot and configure this appropriately. Therefore™ only provides the service that transfers the information from the Therefore™ repository to the customer's Graph connection in their own Azure environment.

What type of information is uploaded?

Once the Graph connection has been configured and the administrator has chosen which categories should be included, the information listed below is uploaded.

- The text of the documents
- The document title
    - Index data is often included in the document title by default. The type of index data available in the document title is fully configurable in Therefore™.
- Ancillary document data such as “Created by”, “Last updated”, etc.
- User access permissions

The following information is NOT uploaded:

- Index data

What types of document formats are supported?

Text-based documents will yield the best results when interacting with Copilot. Uploading documents with images or other non-text-based content is possible, but Copilot will not be able to make effective use of this type of information. To improve the best user experience, we recommend viability testing with the types of documents that users are expected to interact with.

How do users interact with Copilot?

To interact with the data in Therefore™, users must access a Copilot interface (e.g. in a browser, Microsoft Teams, or the Copilot app) and thus require a Copilot license from Microsoft.

The Copilot administrator must make the Graph connection available to users, for example by assigning it to a Copilot Agent and granting users the proper permissions to access this agent.

A user with access to the Graph connection in Copilot can interact with the AI agent and ask questions about information in Therefore™.

How is the user experience determined?

It is important for users to understand that the interactions and experiences they may have with this feature are governed by Microsoft and the Copilot engine. Therefore™ makes the selected information in the repository available to Copilot but has no further influence in how Copilot makes use of the information nor its effectiveness. When using any generative AI engine, users must be aware that answers may be unreliable. There are many factors that will influence a user's experience with this feature, including but not limited to:

- How Copilot or an agent are configured by the administrator (within Azure).
- The description and options configured in the Graph connection.
- The user's own ability to write effective and reliable prompts.
- The other sources of data that may be available to Copilot or an agent.

How are security and permissions handled?

- From a security perspective, users need to be authenticated through Entra ID in both Therefore™ and Copilot. In addition, the Copilot Connector will mirror Therefore™ permissions in full, barring a few exceptions:
    - Permissions are evaluated at the time of document upload. This means the macros 'NOW' and 'TODAY' are not supported and a role will not be granted when using these, with the exception of a a deny role, which will be granted unconditionally.
    - For groups, the macros '$user', '$userno' and '$usermail' are only supported in a basic way. Combinations with 'and', 'or' or, 'in' are not supported.
- When a document is uploaded from Therefore™ to Microsoft Graph, permissions to access this document (Open/View) are evaluated for each user and group. This generates an Access Control List that is checked whenever Copilot requests any information from Microsoft Graph on behalf of a user.
- In order to request information from a document through the Copilot Connector, users will need to:
    - Belong to the same Entra ID tenant as the one used by the Copilot Connector.
    - Be authenticated through Entra ID in both Therefore and Copilot.
    - Have access to said document according to the Access Control List.
- Any internal Therefore™ users that are not in the Entra ID tenant will not be considered by the Connector for the Access Control List.
- The Access Control List is dynamically updated. When changes are made in Therefore™, these will be reflected in the Graph connection. However, due to a Microsoft Graph limitation, it may take between 3 to 6 hours for the permissions to be updated.

Where and how is data stored?

In terms of data security, there are two key areas to address:

- Data in transit: The Therefore™ system passes the information directly from the repository to the customer's own Azure environment in which Copilot is running. The data transfer is encrypted and occurs directly between Therefore™ and Microsoft Azure. The data does not pass through any other servers or third-party service providers.
- Data at rest: While the data is within the Therefore™ Online environment, it is protected like any other data in the repository. The same concept applies to on-premises installations.
- Once uploaded to the customer's Azure environment, the data is stored there and subject to the conditions provided to the customer by Microsoft.

How is data used by Copilot retained in Azure?

Once the data is transferred to the customer's Azure environment, Therefore™ no longer has control of it. Any data retention or security policies within Azure are the responsibility of the Azure environment owner.

Is any data sent from Azure back to Therefore™?

No. This integration only allows one-way transfers of data from Therefore™ to the customer's Azure environment.

How does this integration work with Therefore™ On-Premises?

Please note that this feature will be available for Therefore™ On-Premises starting with Therefore™ 2026.

There are no major differences between using the Copilot integration with Therefore™ Online vs. On-Premises. The server running the service must have access to the internet in order to transfer the data to the customer's Azure environment.

Diagram of content upload from Therefore™ to Microsoft Graph

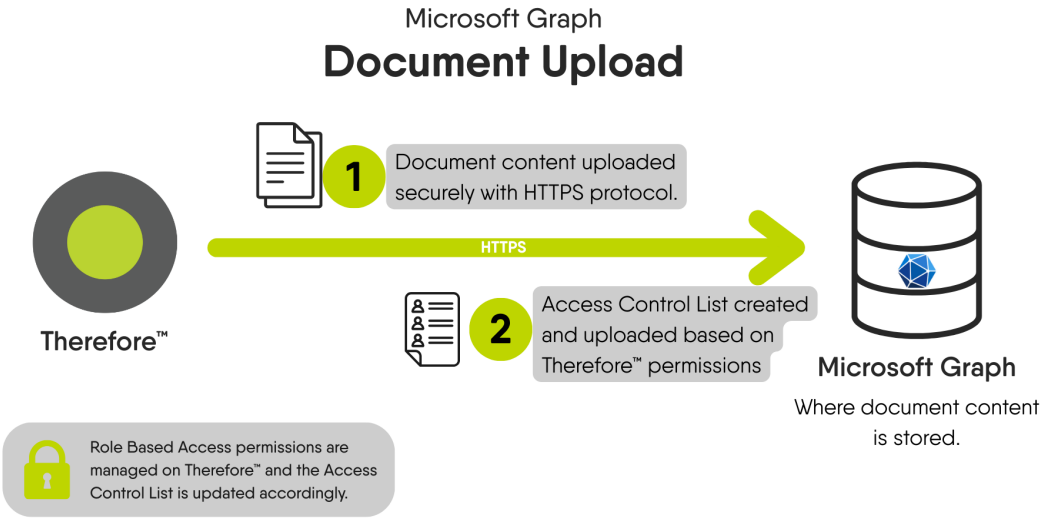
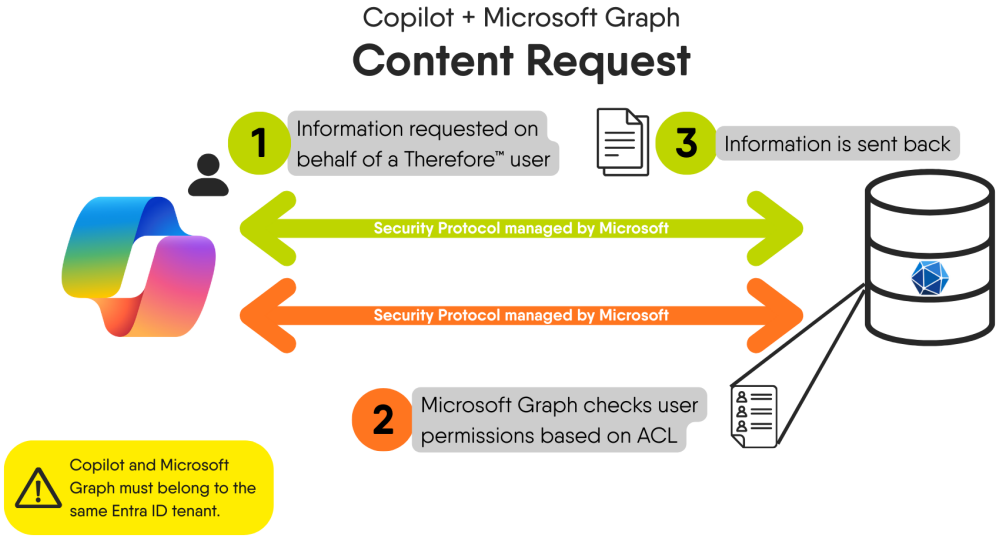


Diagram of Content Request Process



Security Protocol Managed by Microsoft

Communication between Copilot and Microsoft Graph is secured using industry standard protocols and authentication mechanisms managed by Microsoft within the customer's Azure environment. Therefore™ does not control nor have any influence over this protocol.

The following information is sourced from Microsoft.

- **HTTPS/TLS Encryption**
  - All communication between Copilot and Microsoft Graph occurs over HTTPS, secured with Transport Layer Security (TLS).
  - This ensures data in transit is encrypted, preventing eavesdropping or tampering.
- **OAuth 2.0 Authorization**
  - Copilot authenticates to Microsoft Graph using OAuth 2.0.
  - Access tokens are issued by Microsoft Entra ID.
  - Tokens are short-lived and scoped, meaning Copilot only gets the minimum permissions required.
- **OpenID Connect (OIDC)**
  - For identity verification, OIDC (built on OAuth 2.0) is used to validate users and services.
  - This ensures that Copilot is acting on behalf of an authenticated user or service principal.
- **Certificate & Key Management**
  - Microsoft Graph enforces strong certificate validation and secure key exchange during TLS handshakes.
  - This prevents man-in-the-middle attacks.

AI Document Assistant

Is the AI Document Assistant optional?

Yes. The AI Document Assistant is a completely optional feature. It is turned off by default and must be manually enabled by an administrator in the Solution Designer.

Which documents can users interact with using the AI Document Assistant?

The Document Assistant uses only the current document you are viewing and any reference files you deliberately add; it does not search the entire Therefore™ repository. An administrator must select which categories the assistant should be available with. It does not automatically work with every category.

What types of document formats are supported?

Text-based documents will yield the best results. The Document Assistant does not perform OCR or text extraction on any type of document. The exact list of supported formats is:

- PDF (text-based, not image PDF)
- Word, Excel, PowerPoint, OneNote
- Open Office
- Email (MSG, EML)
- Text files (TXT)

How are security and permissions handled?

The feature must first be enabled by an administrator in the Solution Designer, and a mode must be chosen – either Trial or Production mode.

Users must be given at least Read permission on the AI Document Assistant node in the Solution Designer to use the feature.

Otherwise, standard Therefore™ permissions apply. To use the feature, the following conditions must be met:

1. The user must have at least Read permission on the AI Document Assistant node
2. The user must have permission to view/open the document in Therefore™
3. The category where the document is saved must be added to the list of allowed categories in the AI Document Assistant settings

If a user does not have access to view/open a document in Therefore™, they will not be able to use the Document Assistant with it. This also applies to reference documents – a user cannot add a document as a reference to a conversation if they do not have permission to view/open it.

How are conversations retained and when are they deleted?

In Therefore™ v36 (released in November 2026), conversations are only retained for as long as the document is opened in the Therefore™ Viewer or Web Client. This may change in a future version of Therefore™.

To clear the conversation, close the document in the Viewer or Web Client, or press the Reset button in the chat window.

Therefore™ retains the conversation ID and Azure retains the conversation. Conversations are deleted by the server twelve hours after their last use.

How does the Free Trial work?

The Therefore™ AI Document Assistant includes a one-time free trial that allows customers to evaluate the feature before connecting their own Microsoft Foundry account for production use.

The free trial uses Open AI models (e.g. GPT). Upon initial launch (November 2026) GPT 6.0 Luna is used. The exact model in use may be updated periodically by Therefore.

- Eligibility: Every Therefore™ customer is eligible for one trial.
- Activation: The trial is started from the AI Document Assistant configuration in Solution Designer by selecting the trial option and confirming the choice.

- Duration: The trial has no stated time limit. It ends when all allocated tokens have been consumed. The trial cannot be renewed.
- License association: Consumption is tied to the Therefore™ license.
- Trial credits: The number of credits is five times the combined number of named and concurrent user licenses. These credits appear as a license point in Solution Designer.
- Token allocation: Each trial credit provides one million Therefore™ tokens. Changing the number of named or concurrent licenses adjusts the total tokens allocated, however, once a trial is started total amount of tokens will not change even if the number of named or concurrent licenses does.

  Example:

  A Therefore™ Online customer has 5 concurrent licenses and 1 named license. This equals 6 total trial credits, each worth 5 million tokens. In sum, this license will have 30 Million tokens for the free trial.
- Token usage: Trial volume is shared across all users in the system but not restricted to a set amount per user.
- Token calculation: Therefore ™ tokens are calculated from a weighted combination of the input, output, and cached tokens used for each prompt. Output tokens have the highest weighting, while cached tokens have the lowest.
- Estimated capacity: Assuming an average of approximately 8,000 tokens per request, the trial supports roughly 625 requests for each named or concurrent user license. Actual usage varies according to request length and especially the amount of output generated.
- Consumption tracking: Trial-token usage is shown in the AI Document Assistant configuration and in the consumption log. The log records the timestamp, username, document number, and token consumption, but does not store prompts or responses.

- Moving to production: After evaluating the feature, customers can switch from the trial to their own Microsoft Foundry account. When a customer account is used, the trial-token value in the consumption log is shown as zero.

Which AI model is used in Trial mode?

The free trial uses Open AI models (e.g. GPT). Upon initial launch (November 2026) GPT 6.0 Luna is used. The exact model in use may be updated periodically by Therefore.

What is required to move AI Document Assistant into production?

The administrator must simply make the switch from Trial to Production mode in the AI Document Settings in the Solution Designer. This can be done at any time, even if the trial mode was already started and tokens were consumed.

What is logged in the Consumption Log?

The content of conversations (prompts/responses) is not logged. The following data is logged:

- Timestamp
- Username
- Document number
- Token consumption

What information is sent to the AI model when a user asks a question?

- The text of the user prompts during the conversation (whether pre-defined or free-form)
- The text of the currently opened document
- The text of any added referenced documents

Where does AI Document Assistant process document data?

Trial: Data is transmitted to the Azure Foundry of the regional Therefore™ Online system, which uses OpenAI models (e.g. ChatGPT) with a Data Zone deployment wherever possible. The data zone varies by region:

- European Union: The service processes data within the Azure EU Data Boundary. For more details, see: https://learn.microsoft.com/en-us/privacy/eudb/eu-data-boundary-learn
- United States: The service processes data anywhere within the US.
- Asia-Pacific, UK, Canada & Latin America: These regions use a global deployment, meaning the OpenAI models used for the service may be located anywhere in the world (and send the data there). However, data is stored only in the Azure data center where Therefore™ has deployed Azure Foundry – that is, the regional Therefore™ Online system.
    - Data zone configurations are set by Microsoft and cannot be changed by Therefore.

Production (Microsoft Foundry): Customers provision Foundry in their preferred Azure region and use a model of their choice. Thus, this depends on the customer-specific deployment chosen by the Azure environment owner.



Note:

For on-premises systems (starting with v37), the same concept applies. Each on-premises system finds its regional Therefore™ Online system using the license information.

How is data encrypted when using the AI Document Assistant?

Data in transit:

The Therefore™ system passes the information directly from the referenced documents to the AI model. The data transfer is encrypted and occurs directly between Therefore™ and Microsoft Azure. The data does not pass through any other servers or third-party service providers.

Data at rest:

While the data is within the Therefore™ Online environment, it is protected like any other data in the repository. The same concept applies to on-premises installations

- Please note that Therefore™ does not store conversations as of v36. For more details see:
  How are conversations retained and when are they deleted?

Is customer data used to train AI models?

Trial: No.

Production: This depends on the settings of the customer's Azure environment where the Foundry project is hosted. This is under the control and responsibility of the Azure environment owner.

Which third-party AI services or providers are involved in processing the data?

Trial: The data is processed by the OpenAI models (e.g. ChatGPT) provided for the trial within Therefore's Azure environment.

Production: The data is processed in the customer's Azure environment. Any processing or third-party involvement is under the control and responsibility of the Azure environment owner.

How accurate are the answers provided by AI Document Assistant? Can it provide incorrect or misleading answers?

Yes, this is possible. As with any AI model, it is possible for both the model used in the free trial as well as paid models through Microsoft Foundry to provide incorrect or misleading answers. Such instances are not caused by Therefore™, since Therefore™ is simply the document repository. Therefore Corporation encourages responsible use of AI and recommends that humans always check answers provided by any AI model.

Does the quality of the user's prompt affect the response?

Yes. As with any AI model, user prompts have a drastic effect on the responses provided. Therefore Corporation encourages users to write well-defined prompts that include clear instructions about the expected output. To make this easier, the AI Document Assistant allows administrators to:

1. Configure pre-defined prompts instead of relying on free-form prompts by users*.
2. Define general instructions for the AI Document Assistant in the Settings. These are overarching rules that apply to all conversations for all users.

While this makes responses more consistent, it is not infallible. Generative AI engines may produce different responses even with the same prompt and the same reference content.

*Pre-defined prompts can be configured and offered in the conversation window, but the ability for users to write free-form prompts cannot be disabled.

Can AI Document Assistant make decisions or take actions in Therefore™?

No. The Document Assistant is a generative AI tool that can generate text responses to user queries based on document content. It cannot make decisions nor take actions in the Therefore™ system.

Legal

All rights reserved. No parts of this work may be reproduced in any form or by any means - graphic, electronic, or mechanical, including photocopying, recording, taping, or information storage and retrieval systems - without the written permission of the publisher.

3rd Party Products that are referred to in this document, are the property of, and may be either trademarks and/or registered trademarks of the respective owners in the USA and/or other countries. The publisher and the author make no claim to these trademarks.

While care has been taken in the preparation of this document, the publisher and the author assume no responsibility for errors or omissions, or for damages resulting from the use of information contained in this document.