

ASUS AI SuperBuild

User Guide

Table of Contents

Chapter 1: Getting Started	2
1.1 System Requirements.....	2
1.2 Installation	3
1.3 System Information & Updates.....	14
Chapter 2: The AI SuperBuild Interface	17
2.1 Main Chat Interface	17
2.2 Managing Chat History.....	18
Chapter 3: Assistant Configuration	19
3.1 AI Assistant List	19
3.2 Model Upload and Convert.....	20
3.3 Creating a New Assistant	28
3.4 Editing an Existing Assistant	34
3.5 Knowledge Base Management.....	34
3.6 Evaluation and Report.....	35
3.7 One-Key Import/Export	51
Chapter 4: MCP (Model-Context-Protocol).....	55
4.1 Example: Build a Hotel Assistant MCP Server	55
4.2 MCP Marketplace.....	61
Chapter 5: Vision Language Model(Image Analysis)	65
5.1 Executing VLM in the Main Chat	65
Appendix	66
1. MCP Example Setting SOP	66

Chapter 1: Getting Started

This chapter provides an overview of the system requirements and the installation process for AI SuperBuild.

1.1 System Requirements

Before installing AI SuperBuild, please ensure your system meets the minimum hardware and software requirements outlined below. Meeting the recommended requirements will provide a better user experience, especially when working with multiple or larger language models.

Hardware Requirements

Component	Minimum Requirements	Recommended Requirements
Processor	Intel® Core™ Ultra processor Series 1 (Meteor Lake)	Intel® Core™ Ultra 200V series (Lunar Lake)
Memory (RAM)	16GB	32GB
Storage	4GB for AI Assistant with 1 LLM	12GB for AI Assistant with 3 LLMs
Graphics	Integrated Intel® Graphics	Integrated Intel® Arc™ Graphics
Network	Broadband connection for LLMs and component downloads	

Note:

- AI SuperBuild has been validated on limited Intel AIPC: NUC 14 Pro, NUC 14 Pro AI, NUC 14 Pro AI+, NUC 15 Pro and NUC 16 Pro.
- Minimum Intel Graphics driver version is 30.0.100.9955, and the minimum NPU driver version is 32.0.100.3714. [Please visit the Intel Download Center for the latest drivers](#)

Software Requirements

Microsoft Windows 11 (Version 23H2 or newer) is required. During installation, the AI SuperBuild application may download and install additional required components.

1.2 Installation

The following steps will guide you through the installation of AI SuperBuild. The process includes the setup of required prerequisite software, followed by the main application installation and first-time model download.

Installation Steps


1. Unzip the Package:

Begin by unzipping the downloaded ASUS_AI_SuperBuild_x.x.x.



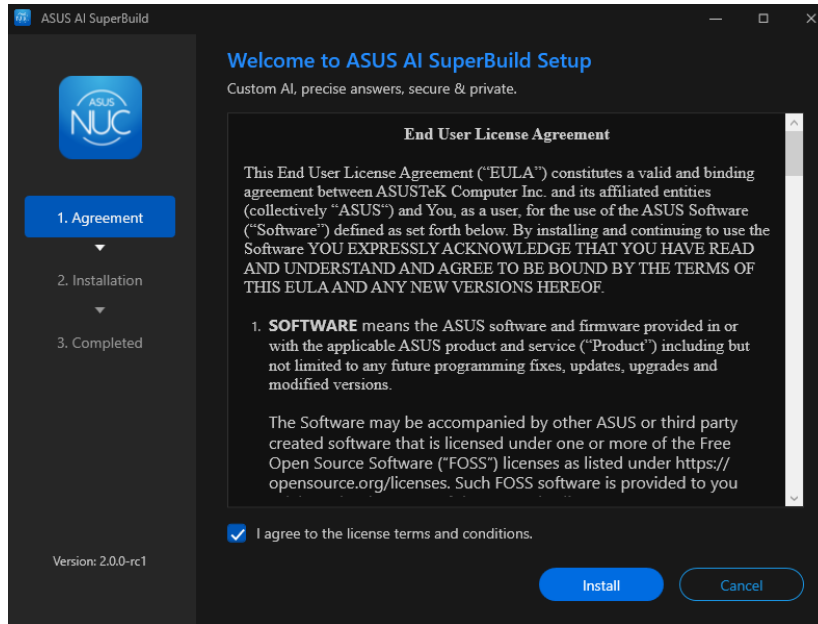
2. Run the Installer Script:

Open the extracted folder and double-click “AISuperBuild_Installer”. This installer will initiate the entire installation process.

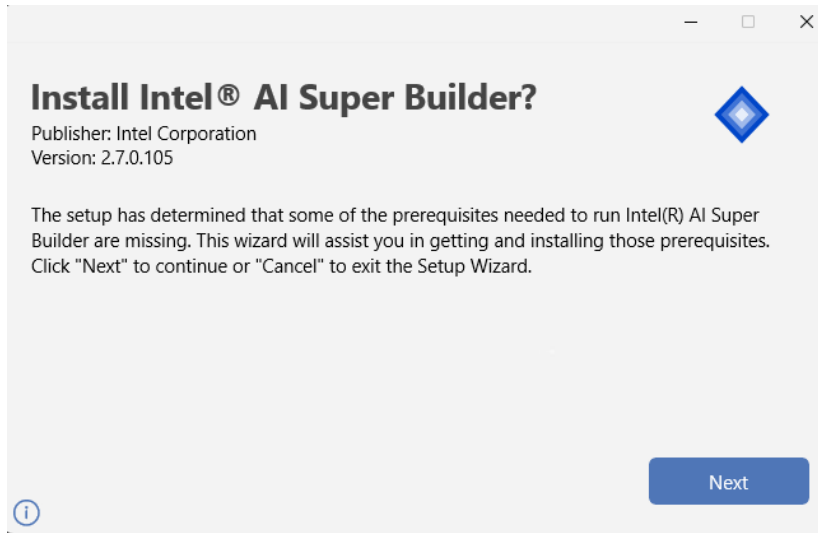
Name	Date modified	Type
 AISuperBuild_Installer	09/02/2026 15:07	Application

3. Follow the On-Screen Prompts:

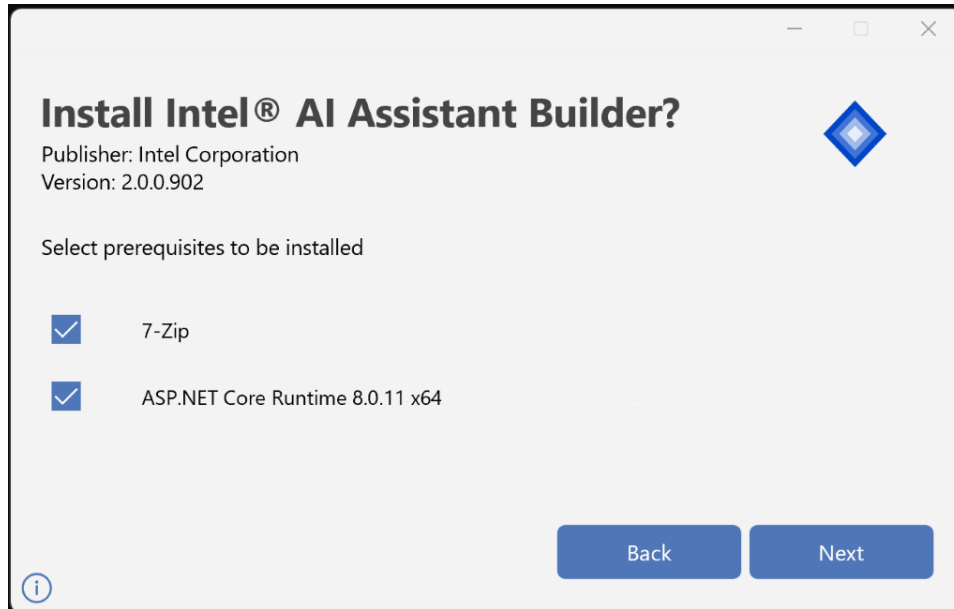
A series of installation wizards will appear. Follow the on-screen instructions, clicking Next, Install, and Accept as prompted to install the prerequisite software (Intel® AI Assistant Builder, 7-Zip, ASP.NET Core Runtime) and the main AI SuperBuild application.



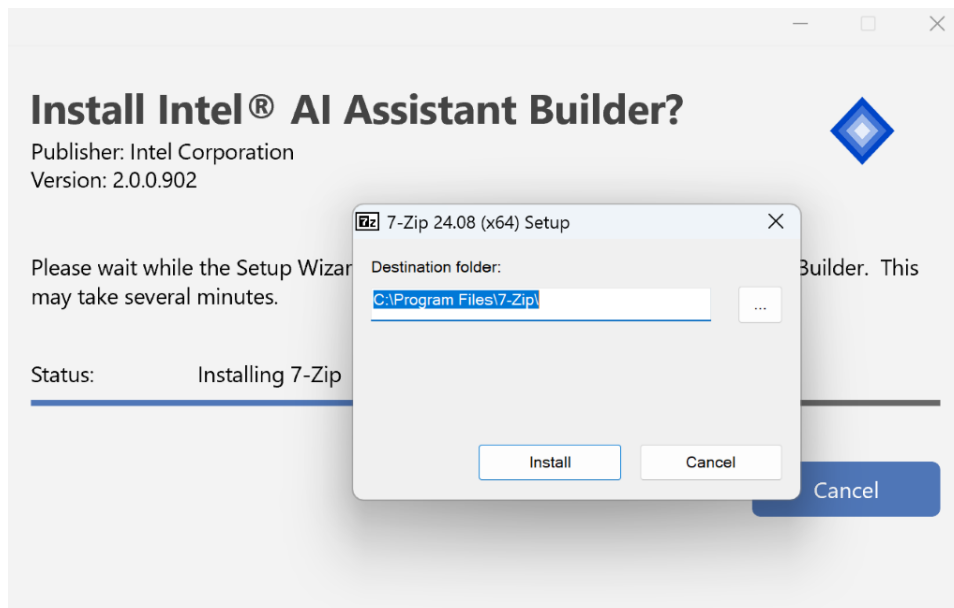
Select to agree to the license terms and conditions and select install



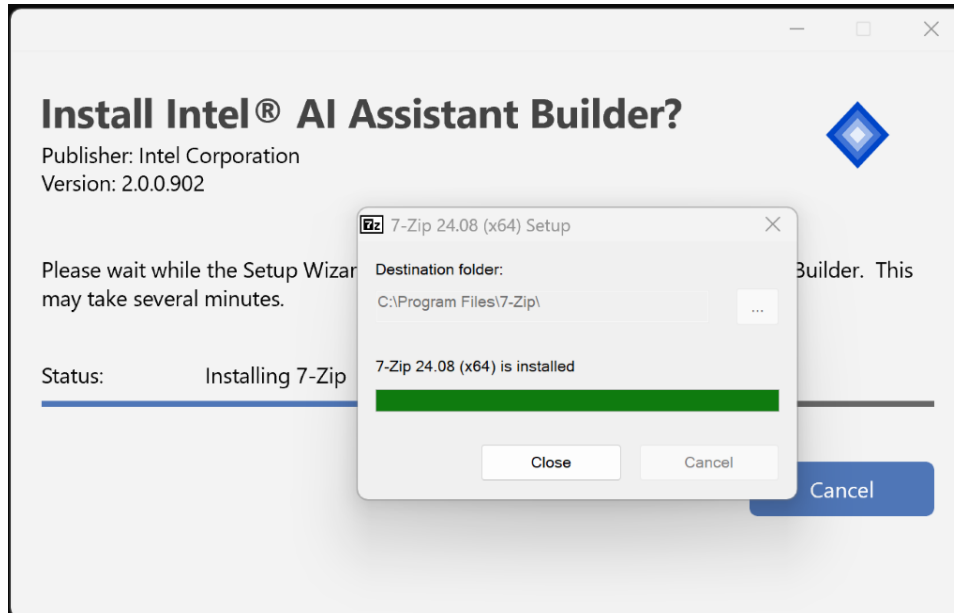
Select Next



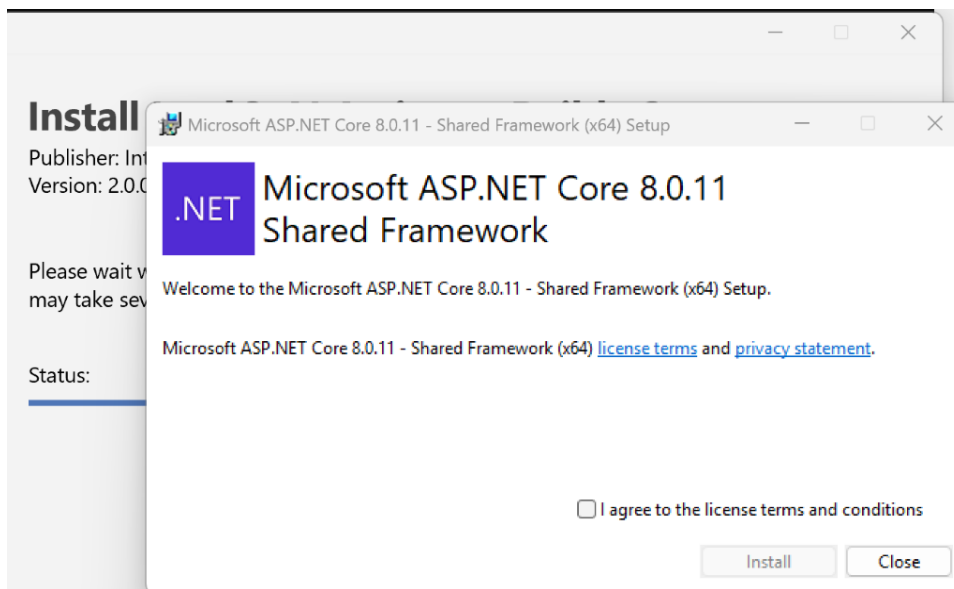
Select Next



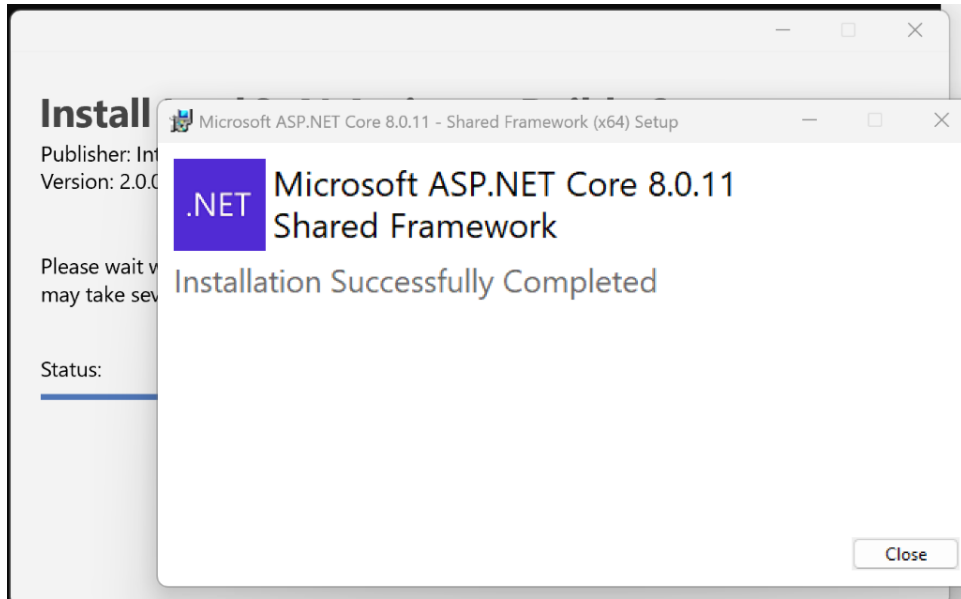
Select Install



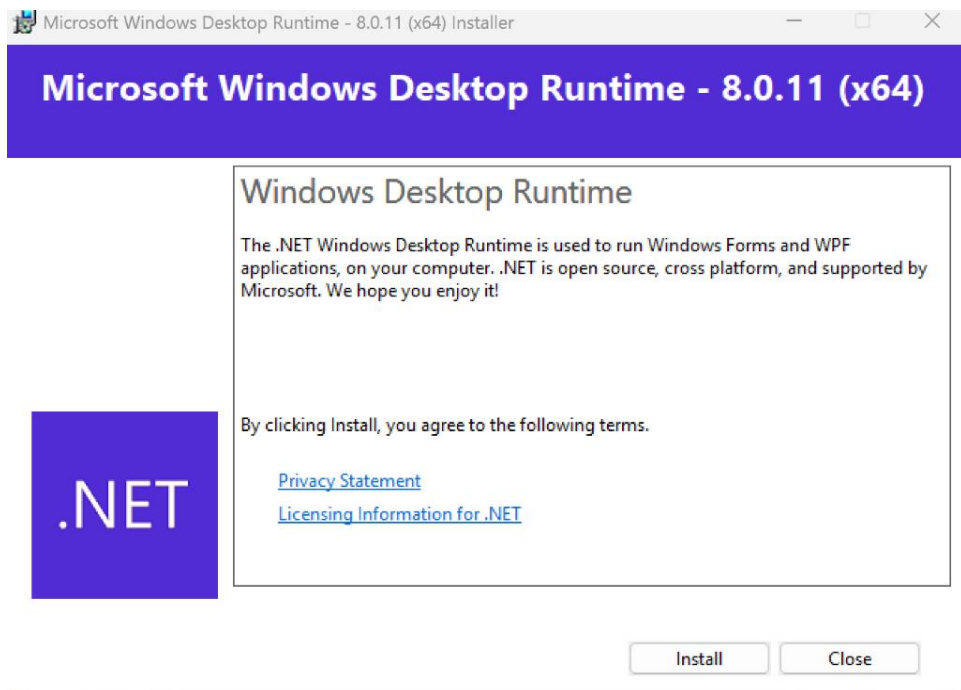
Select Close



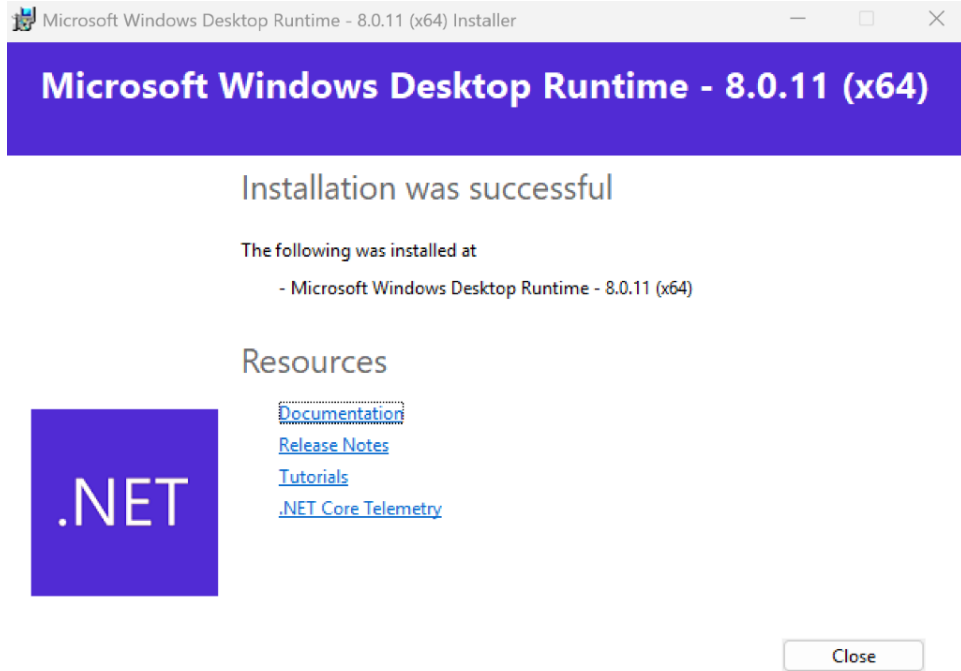
I agree, then install



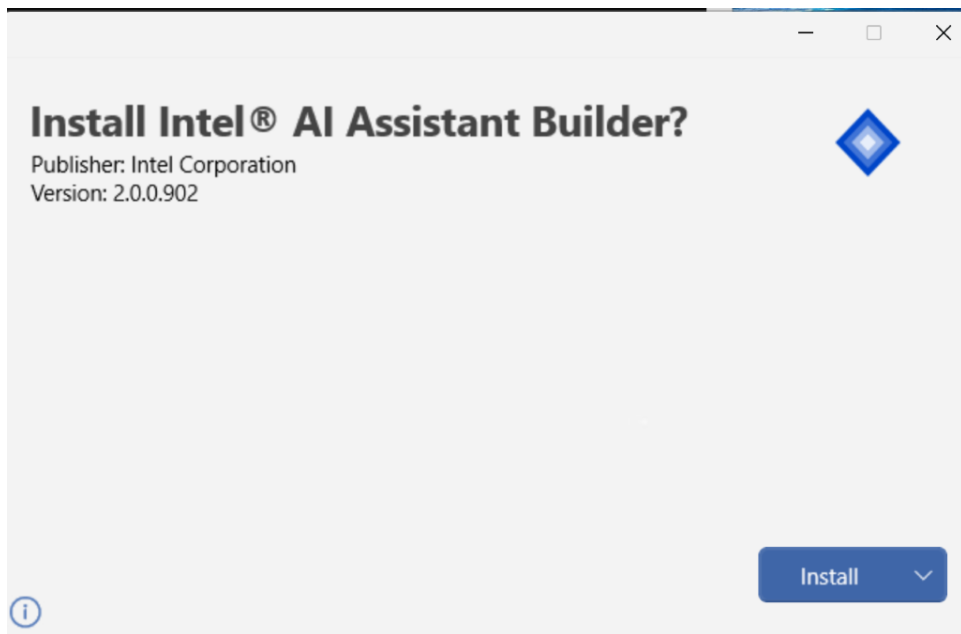
Select Close



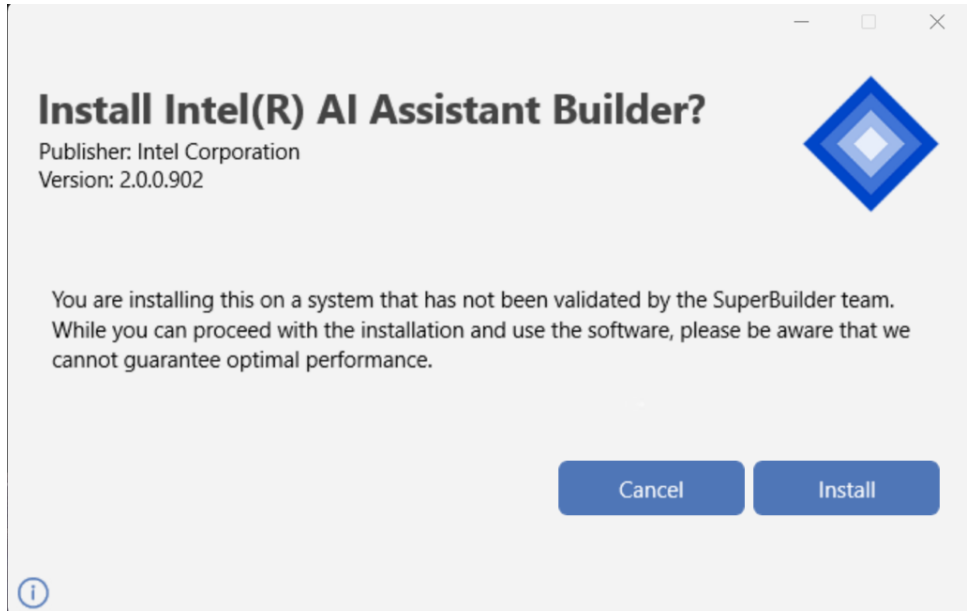
Select Install



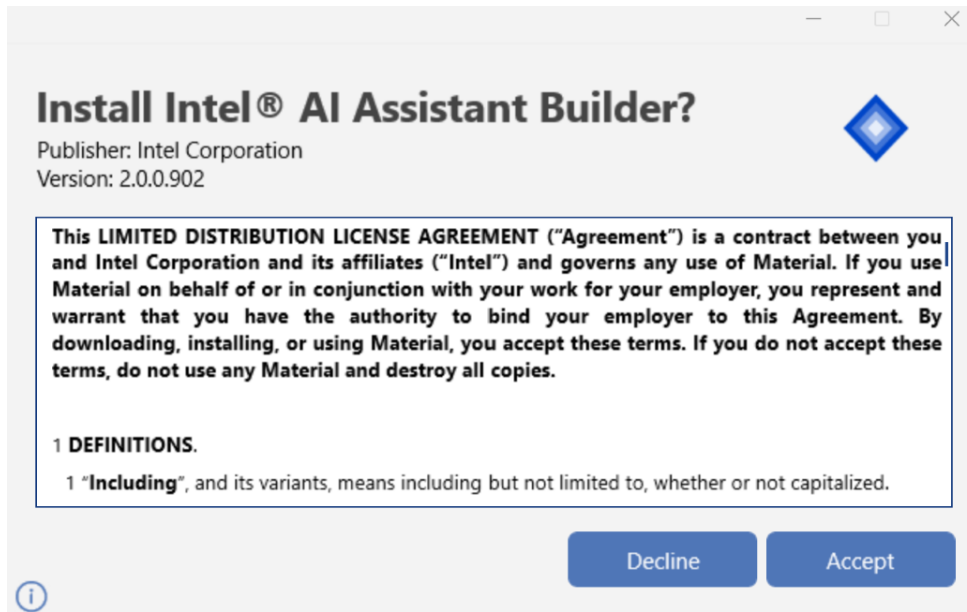
Select Close



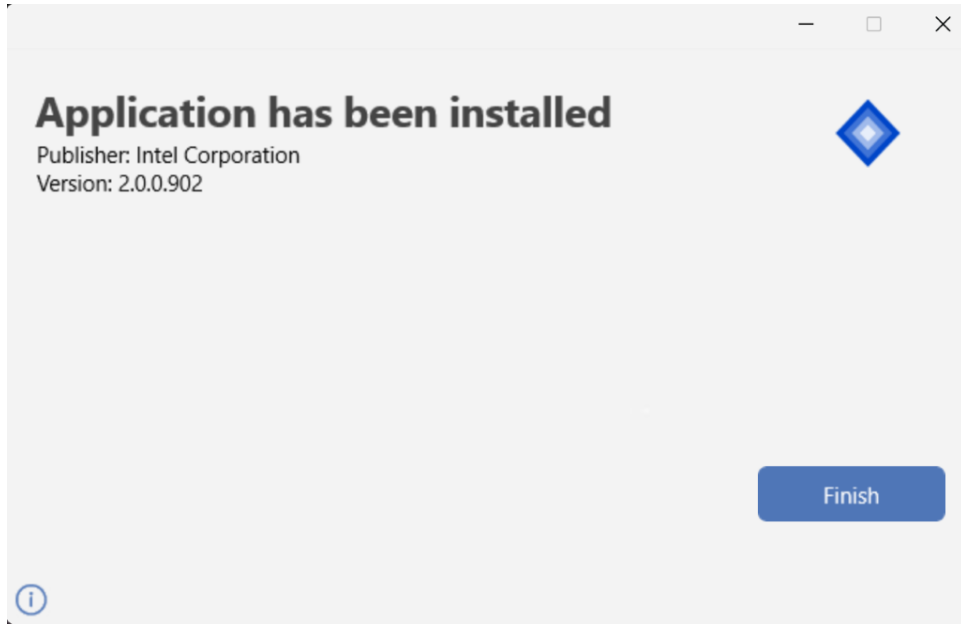
Install



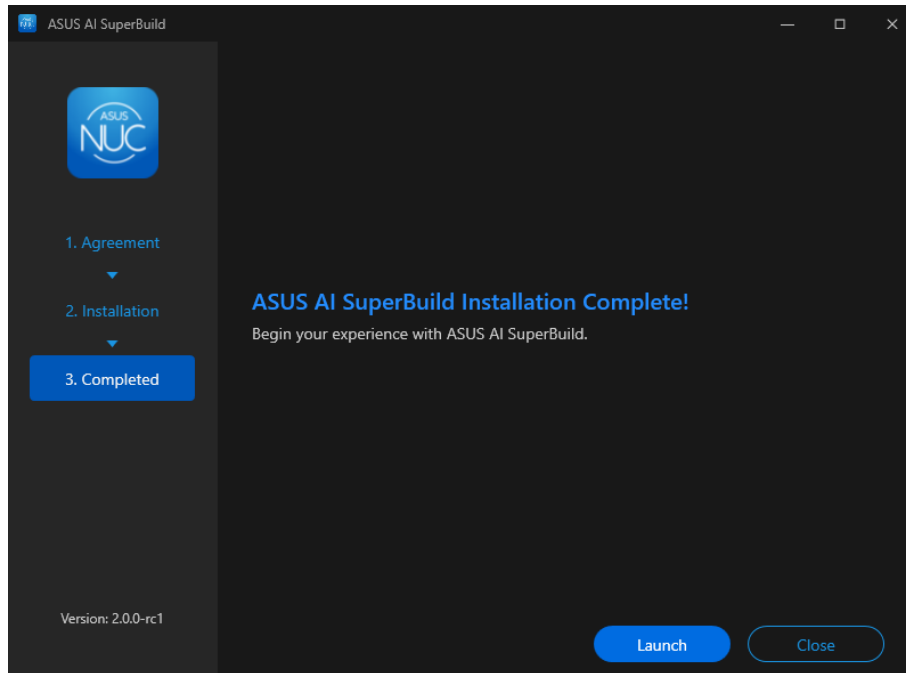
Accept



Finish



Select Finish



Installation Complete

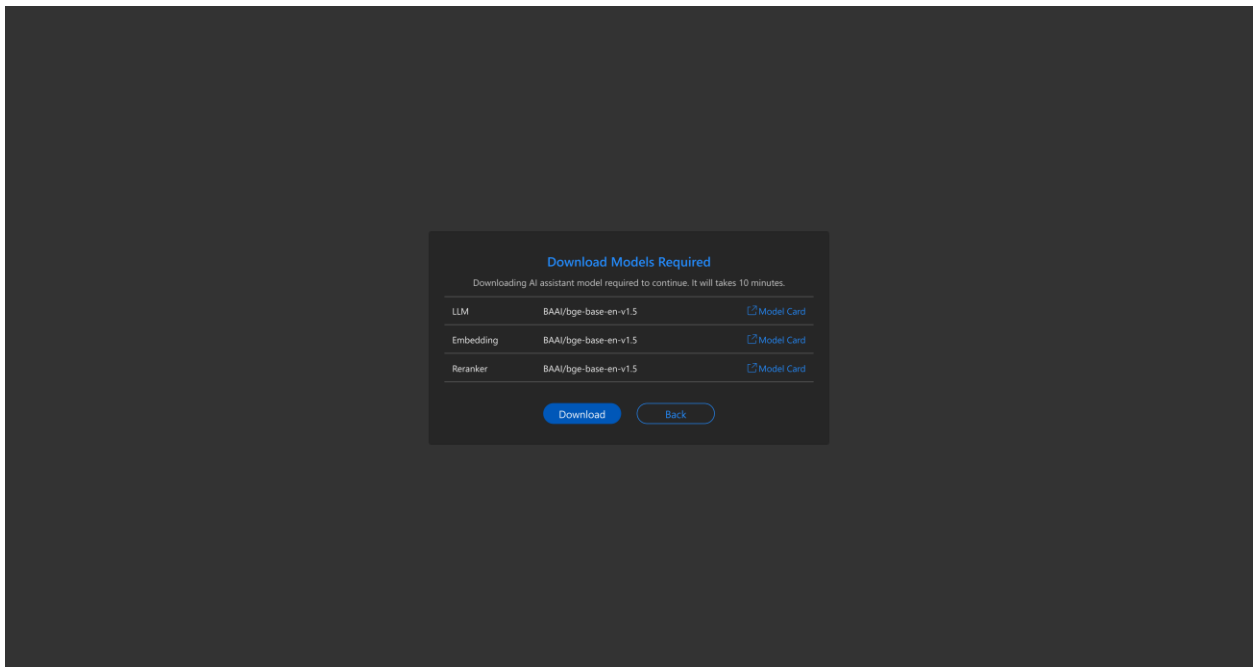
4. Launch the Application:
Once the installation is complete if you want to launch ASUS AI SuperBuild. Click Launch. A new desktop shortcut will also be created for future access.



First-Time Setup: Model Download

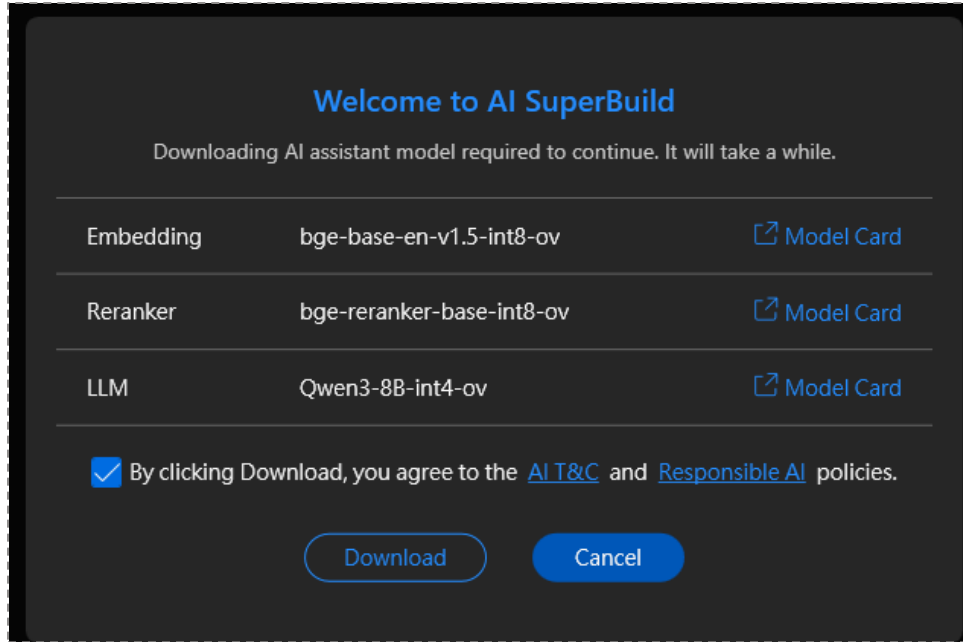
1. Initial Model Download:

Upon the first launch, a "Welcome to AI SuperBuild" window will appear, prompting you to download the required AI assistant models (Embedding, Reranker, and LLM).



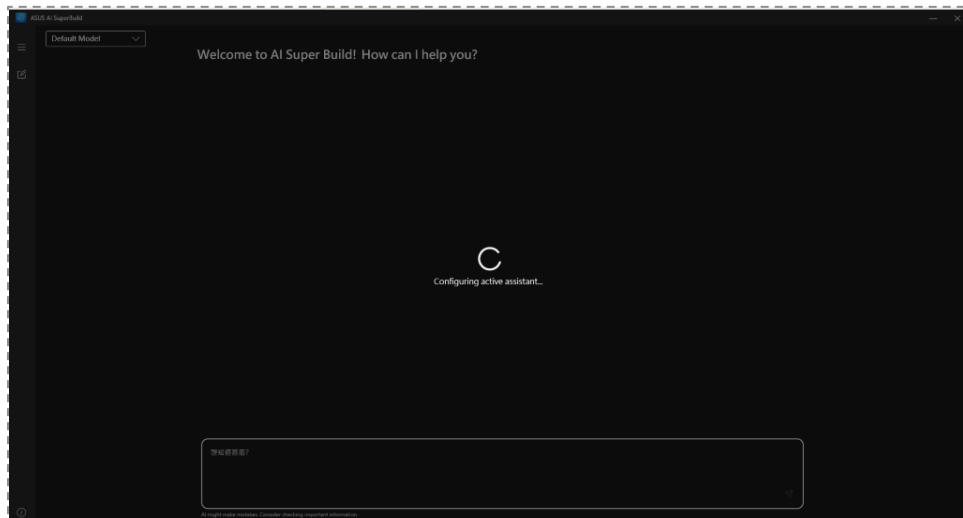
2. Agree and Download:

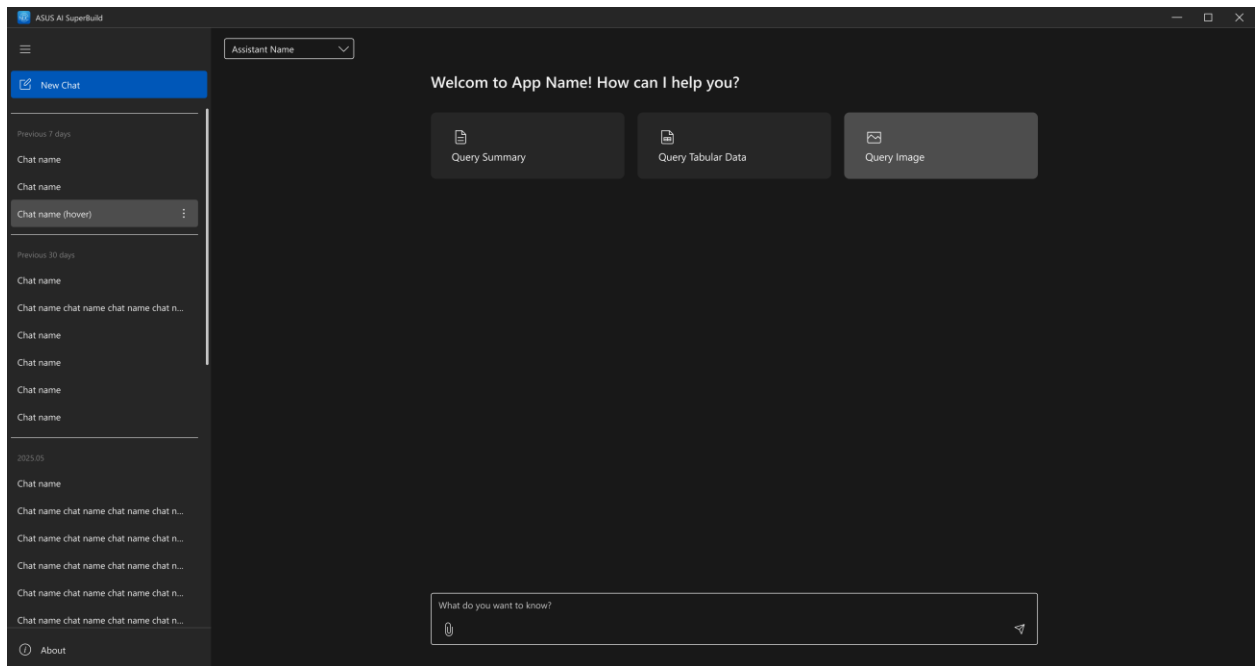
Check the box to agree to the AI T&C and Responsible AI policies, then click Download. The process may take around 15 minutes, depending on your internet speed.



3. Configuration and Use:

After the download completes, the application will configure the active assistant. You can then begin using your AI SuperBuild assistant.





Known Issues

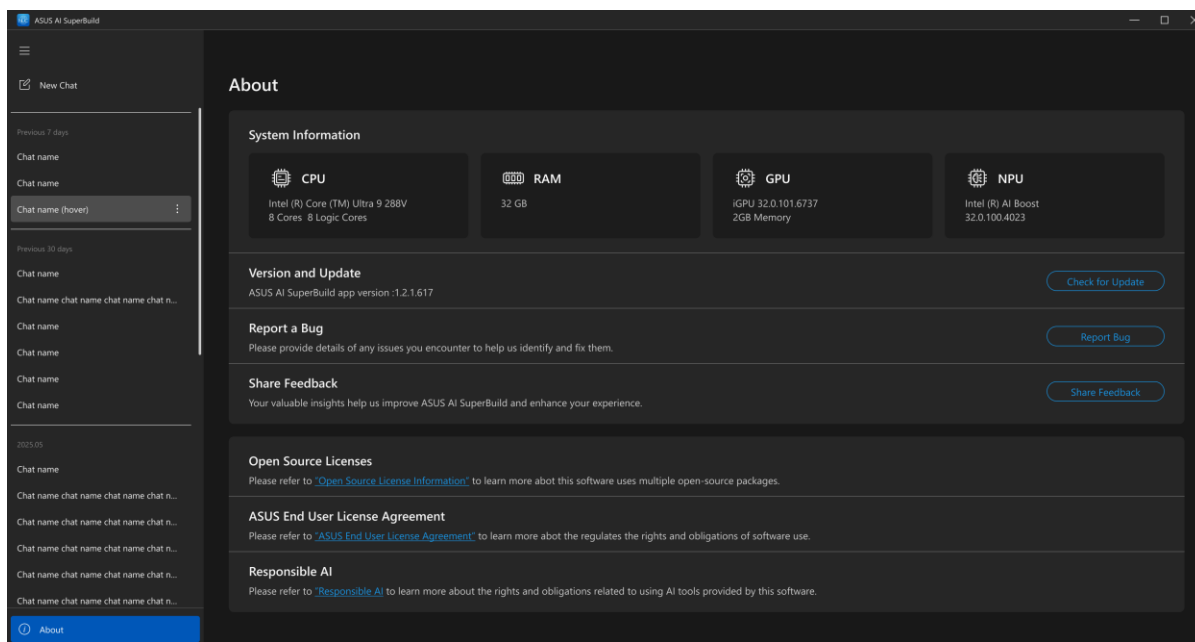
- When the software is used for the first time, it may take approximately 35 to 50 seconds to respond to the initial prompt.
- Text generation speed may be slow in the current version. This is slated for improvement in a future release.

1.3 System Information & Updates

This chapter explains how to access system information, check for software updates, and find legal and compliance information related to AI SuperBuild.

1.3.1 About Page

The "About" page provides a comprehensive overview of your system's hardware, the AI SuperBuild software version, and links to important documentation. It is the central place to verify your setup and ensure your software is up to date.



[Check for Update](#)

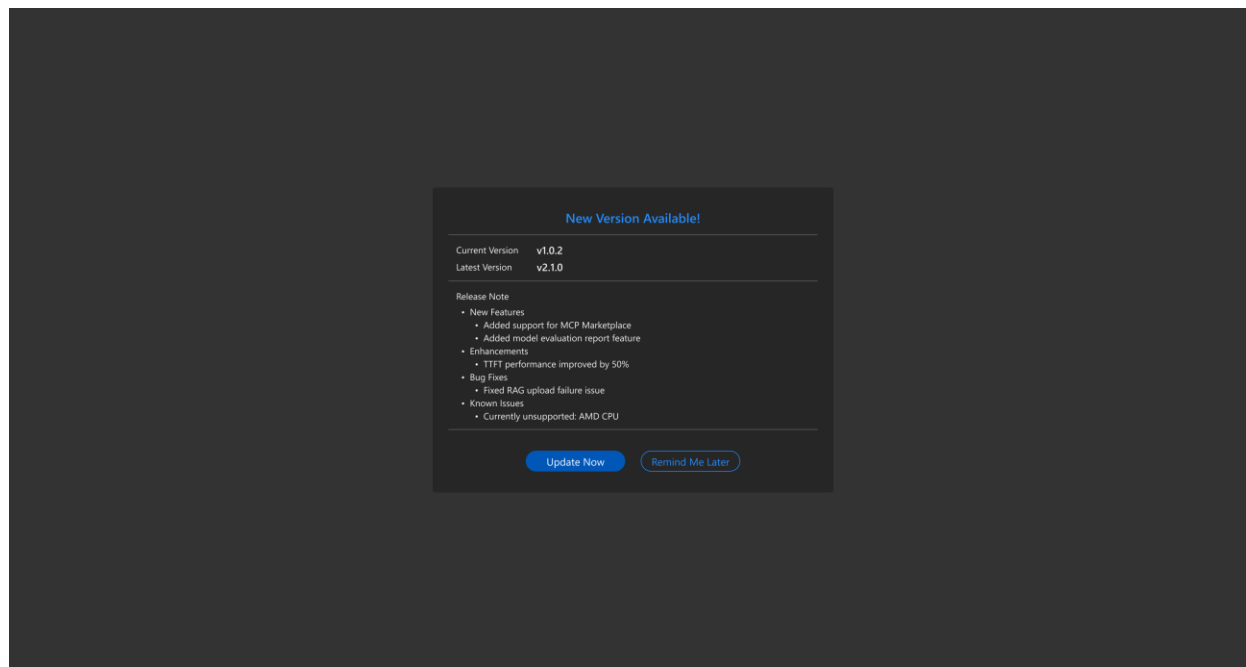
UI Element	Description
System Information	This section displays the key hardware components of your system.
CPU	Shows the processor model, number of cores, and logical processors. (e.g., Intel(R) Core(TM) Ultra 9 288V)
RAM	Shows the total amount of installed system memory (e.g., 32 GB).
GPU	Displays the graphics processing unit model, driver version, and memory (e.g.,

	GPU 32.0.101.6737, 2GB Memory).
NPU	Displays the Neural Processing Unit model and driver version, which is used for accelerating AI tasks (e.g., Intel(R) AI Boost 32.0.100.4023).
Version and Update	This section shows the currently installed application version of AI SuperBuild (e.g., 1.1.0)
Check for Update	Clicks to check for and download the latest version of the application.
Report a bug	Clicking this option will open a Google Form, allowing you to report any bugs or system issues directly to ASUS.
Share Feedback	We value your input! Click this button to send any feedback, ideas, or comments regarding your AI SuperBuild experience directly to the ASUS team.
Open Source Licenses	Opens a new window with information about the open-source software packages used in the application.
ASUS End User License Agreement	Opens the EULA, which details the rights and obligations of software use.
AI T&C	Opens the Terms and Conditions related to using the AI functions within the software.
Responsible AI	Opens documentation regarding the responsible and ethical use of the AI tools provided by the software.

1.3.2 Check for update

Each time you launch the AI SuperBuild, the system automatically checks for the latest software version. If an update is available, simply click Update Now, and the application will handle the download and installation automatically.

Alternatively, you can manually check for new versions at any time by navigating to the Settings page and clicking Check for update.

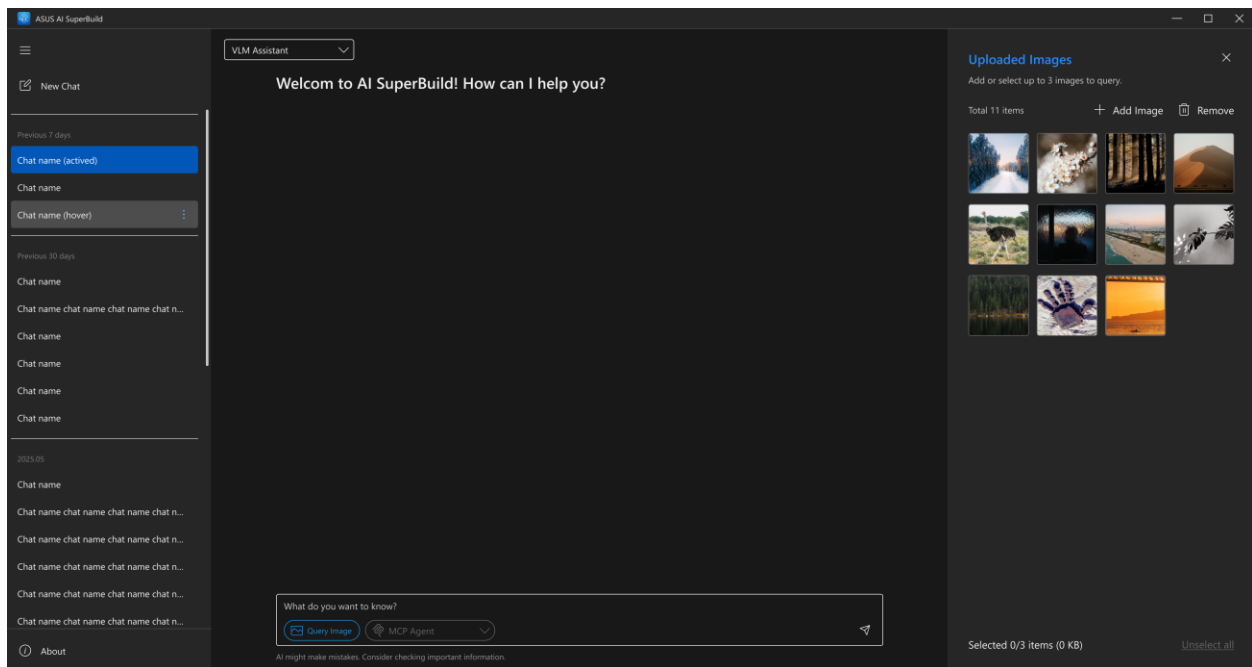


Chapter 2: The AI SuperBuild Interface

This chapter describes the main user interface and core functionalities for interacting with the AI Assistant, including how to start chats, select models, and manage your conversation history.

2.1 Main Chat Interface

The main interface is your central hub for interacting with the AI SuperBuild assistant. The interface features an intuitive layout with a left-hand navigation panel for managing chats, a central conversation window, and input tools at the bottom.

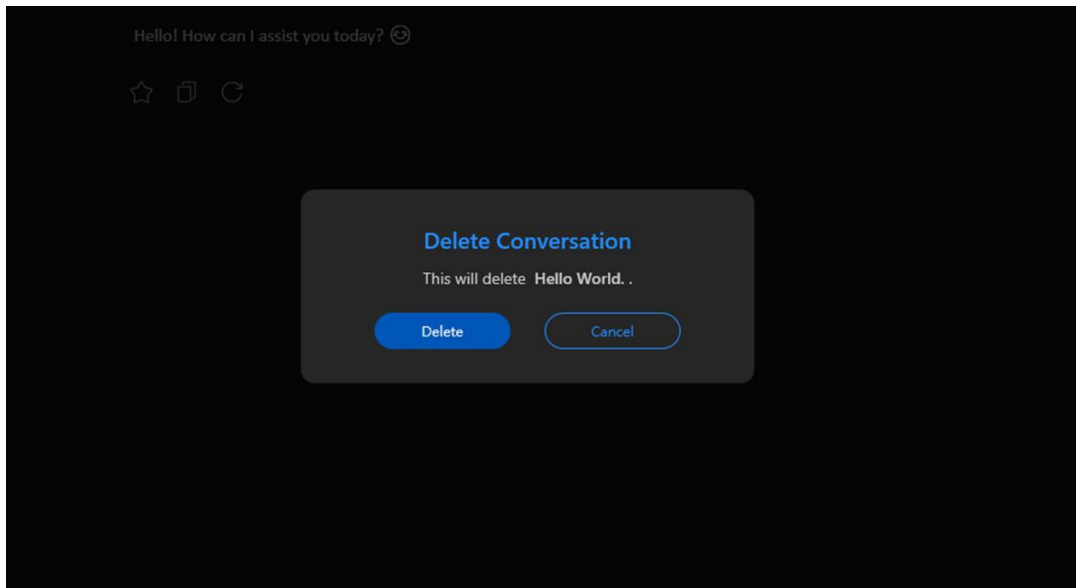
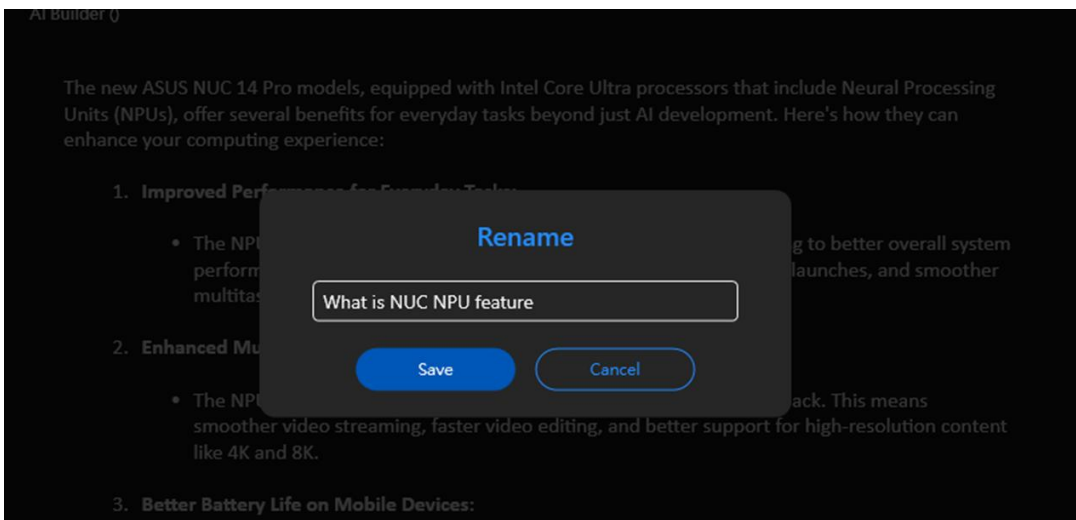


UI Element	Description
Menu (☰)	Toggles the visibility of the left-hand navigation panel.
New Chat	Submits your text prompt to the AI assistant.
Query images	Upload any image or screenshot for automatic content analysis. The AI can instantly interpret charts, analyze user interface (UI) layouts, and extract or summarize text from document images.
Model Selection	Allows you to select the active AI model for your session from a list of available models like "Default Model", "Model A", etc.

Model Configuration	Navigates to the advanced settings page to create, edit, and manage your AI assistants.
Text Input Field	The main area where you type your questions or prompts for the AI assistant.
Send Icon	Clicks to submit your text prompt to the AI assistant for a response.

2.2 Managing Chat History

The left-hand navigation panel displays your conversation history, organized by date. You can easily revisit, rename, or delete past conversations.



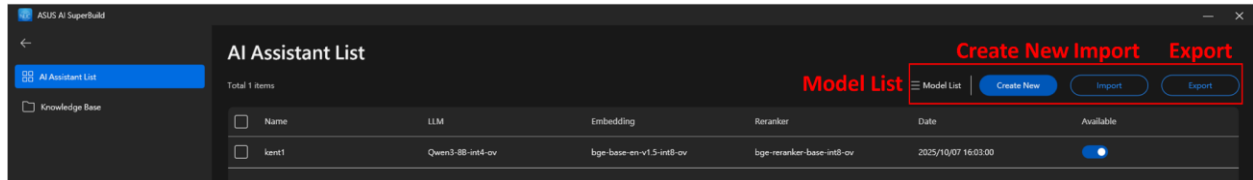
Action	Description
Select a Chat	Click on any chat name in the history to load it into the main window and continue the conversation.
Rename a Chat	Right-click on a chat and select "Rename" to give it a more descriptive name for easier identification.
Delete a Chat	Right-click on a chat and select "Delete" to permanently remove the conversation.

Chapter 3: Assistant Configuration

This chapter covers the advanced settings for creating and managing your AI Assistants. You can access this section by navigating from the **Model dropdown list > Model Configuration**.

3.1 AI Assistant List

This page displays the configuration settings for all assistants. Each assistant is defined by a specific Large Language Model (LLM) and its associated parameters for embedding and reranking, which determines how it processes and responds to queries.



Button/UI Element	Description
Create New	Opens the "New Assistant" wizard to create a new, customized AI assistant profile from scratch.
Import	Allows you to import a previously exported assistant configuration file, making it easy to share or restore settings.
Export	Allows you to export the selected assistant's configuration as a file for backup or deployment on another machine.
Model List	Navigates to a screen where you can manage all the AI models (LLM, Embedding, Reranker) available in the system.

Available (Toggle)	Toggles the assistant's availability in the main chat Model Selection dropdown.
---------------------------	---

3.2 Model Upload and Convert

This chapter explains how to upload and convert models in AI SuperBuild. This feature allows you to import models from local storage or third-party platforms like **Hugging Face** and **ModelScope**. This allows you to extend beyond the default 8 built-in LLM models.

AI SuperBuild currently supports the following:

- **Model Sites**
 - Hugging Face
 - ModelScope
 - Local Path (self-built or locally prepared models)
- **Model types**
 - LLM (Large Language Model)
 - Embedding model
 - Reranker model

Note

Currently, AI SuperBuild only supports **text generation LLM models**. Make sure the model you choose is a text generation model.

3.2.1 Model Upload

If you already have models, you can make them available to AI SuperBuild in two ways:

- **Upload models by folder**
 - In AI SuperBuild, navigate to **AI Assistant List** → **Model List** → **ADD** → **Upload**.
 - **Copy** or **move** the entire model folder into the designated AI SuperBuild model directory.
 - This is recommended when you have already downloaded models (for example from Hugging Face, ModelScope, or other sources) and just want AI SuperBuild to recognize them.
 - Follow the on-screen instructions to select and upload your prepared model folder.

Upload Model

The model must be properly converted to run with the NUC AI Super Build application.
The procedure provided may not work with all models.

Model Type

LLM
▼

Upload Method

Copy Folder
▼

Local Folder

Browse...

Upload

Cancel

3.2.2 Model Convert

To download, convert models from Hugging Face, ModelScope, or a local path, use the **Model Convert** feature.

1. Navigate to **AI Assistant List** → **Model List** → **ADD** → **Convert**.
2. You will see:
 - A selector for **Model Site** (Hugging Face / ModelScope / Local Path)
 - A selector for **Model Type** (LLM / Embedding / Reranker)
 - An input field for **Model ID** or local path
 - A **Convert** button and related status information

ASUS AI SuperBuild

AI Assistant List > Model List

Model List

LLM Embedding Reranker

Total 9 items

Name	Description	Source	Status
Qwen2.5-7B-Instruct-int4-ov	qwen2.5	https://huggingface.co/OpenVINO/Qwen2.5-7B-Instruct-int4-ov	✓ Available
Qwen3-8B-int4-ov	qwen3	https://huggingface.co/OpenVINO/Qwen3-8B-int4-ov	✓ Available
Phi-4-mini-instruct-int4-ov	phi4-mini	https://huggingface.co/OpenVINO/Phi-4-mini-instruct-int4-ov	✓ Available
Phi-3.5-vision-instruct-int4-ov	Phi-3.5-vision-ov	https://huggingface.co/OpenVINO/Phi-3.5-vision-instruct-int4-ov	✓ Available
DeepSeek-R1-Distill-Qwen-7B-int4-ov	deepseek-r1-qwen	https://huggingface.co/OpenVINO/DeepSeek-R1-Distill-Qwen-7B-int4-ov	✓ Available
zephyr-7b-beta-int4-ov	zephyr7b	https://huggingface.co/OpenVINO/zephyr-7b-beta-int4-ov	Download
BioMistral-7B-SLERP-int4	biomistral7b-slerp	https://ai.builder.intel.com/models/BioMistral-7B-SLERP-int4.zip	Download
Phi-3.5-mini-instruct-int4-cw-ov-npu	Phi-3.5-mini-npu	https://huggingface.co/OpenVINO/Phi-3.5-mini-instruct-int4-cw-ov	Download
Nemotron-Cascade-8B	Custom Uploaded Mox	https://huggingface.co/mudra/Nemotron-Cascade-8B	✓ Available

Convert Model

Select a model from huggingface, modelscope, or your local machine to convert it into an OpenVINO-compatible artifacts.

Model Site

Model Type

Model ID ⓘ

Hugging Face Token (Optional)

3.2.3 Model Convert

To download, convert models from Hugging Face, ModelScope, or a local path, use the **Model Convert** feature.

1. Navigate to **AI Assistant List** → **Model List** → **ADD** → **Convert**.
2. You will see:
 - A selector for **Model Site** (Hugging Face / ModelScope / Local Path)
 - A selector for **Model Type** (LLM / Embedding / Reranker)
 - An input field for **Model ID** or local path
 - A **Convert** button and related status information

ASUS AI SuperBuild

AI Assistant List > Model List

Model List

LLM Embedding Reranker

Total 9 items

Name	Description	Source	Status
Qwen2.5-7B-Instruct-int4-ov	qwen2.5	https://huggingface.co/OpenVINO/Qwen2.5-7B-Instruct-int4-ov	✓ Available
Qwen3-8B-int4-ov	qwen3	https://huggingface.co/OpenVINO/Qwen3-8B-int4-ov	✓ Available
Phi-4-mini-instruct-int4-ov	phi4-mini	https://huggingface.co/OpenVINO/Phi-4-mini-instruct-int4-ov	✓ Available
Phi-3.5-vision-instruct-int4-ov	Phi-3.5-vision-ov	https://huggingface.co/OpenVINO/Phi-3.5-vision-instruct-int4-ov	✓ Available
DeepSeek-R1-Distill-Qwen-7B-int4-ov	deepseek-r1-qwen	https://huggingface.co/OpenVINO/DeepSeek-R1-Distill-Qwen-7B-int4-ov	✓ Available
zephyr-7b-beta-int4-ov	zephyr7b	https://huggingface.co/OpenVINO/zephyr-7b-beta-int4-ov	Download
BioMistral-7B-SLERP-int4	biomistral7b-slerp	https://ai.builder.intel.com/models/BioMistral-7B-SLERP-int4.zip	Download
Phi-3.5-mini-instruct-int4-cw-ov-npu	Phi-3.5-mini-npu	https://huggingface.co/OpenVINO/Phi-3.5-mini-instruct-int4-cw-ov	Download
Nemotron-Cascade-8B	Custom Uploaded Mox	https://huggingface.co/mudra/Nemotron-Cascade-8B	✓ Available

In the **Model Site** section, choose where you want to get the model from:

- **Hugging Face** – Download and convert a model directly from the Hugging Face Hub.
- **ModelScope** - Use a model hosted on the ModelScope platform.
- **Local Path** – Use a model that has already been prepared on your local environment.

After you select a source, the corresponding input field (**Model ID** or **local path**) will be enabled.

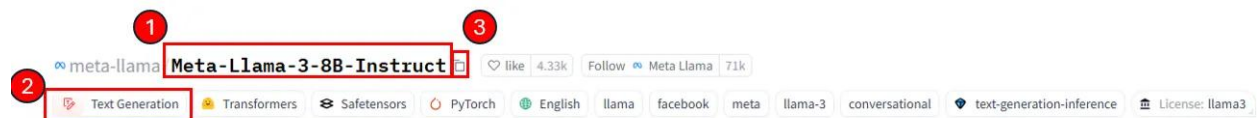
Using Hugging Face

If you choose **Hugging Face** as the model source, follow these steps to obtain the correct model ID:

1. Open the Hugging Face website:

huggingface.co

2. Use the search bar to find the model you want to use (for example: meta-llama/Meta-Llama-3-8B-Instruct).
3. In the search results, click the target model to open its detail page.
4. On the model page, click the **Copy** button next to the model name to copy the full **model ID**.



Important

Confirm that the model is a **Text Generation** model so that it can work correctly in AI SuperBuild.

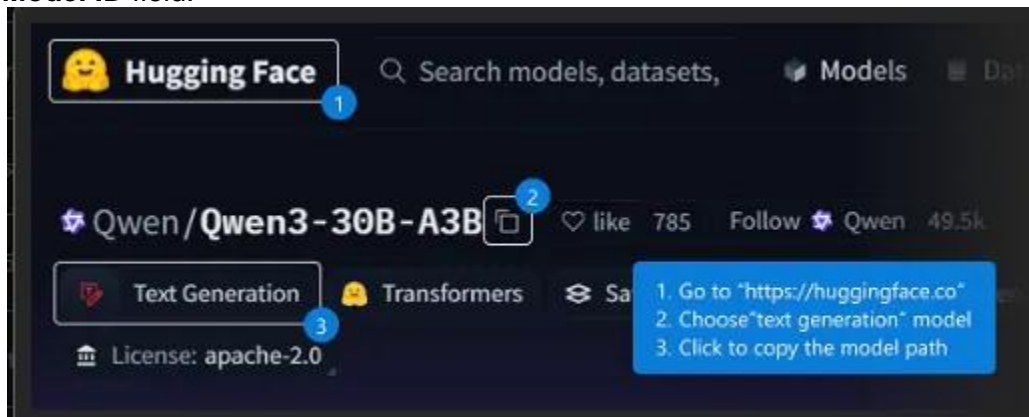
Using ModelScope

If you choose **ModelScope** as the model source:

- Go to the ModelScope website and locate the target model.
- Copy the **model ID** from the ModelScope page (similar to Hugging Face).
- Confirm that the model type and task are supported by AI SuperBuild (text generation LLM, embedding, or reranker).
- Paste the ModelScope model ID into the **Model ID** field in AI SuperBuild.

Note

The overall process is the same as Hugging Face: copy the model ID from ModelScope and use it in the **Model ID** field.

**Using Local Path models**

If you choose **Local Path** as the model source:

- Prepare the model files in a local directory.
- Set **Model Site** to **Local Path** and specify the local directory path in the **Model ID / Path** field.

Tip

This is recommended when your models are hosted on an internal file server or have already been converted by your ML/infra team.

Convert Model

Select a model from huggingface, modelscope, or your local machine to convert it into an OpenVINO-compatible artifacts.

Model Site
Local Path

Model Type
LLM

Local Folder
Browse...

Convert Cancel

Select the Model Type

In the **Model Type** section, choose the type of model you want to convert:

- **LLM** – For text generation language models.
- **Embedding** – For models that generate vector embeddings from text.
- **Reranker** – For models used to re-rank search or retrieval results.

Make sure the selected model type matches the actual purpose and architecture of the model. For example, text generation models should be configured as **LLM**.

Convert Model

Select a model from huggingface, modelscope, or your local machine to convert it into an OpenVINO-compatible artifacts.

Model Site
Hugging Face

Model ID ⓘ

Hugging Face Token (Optional)

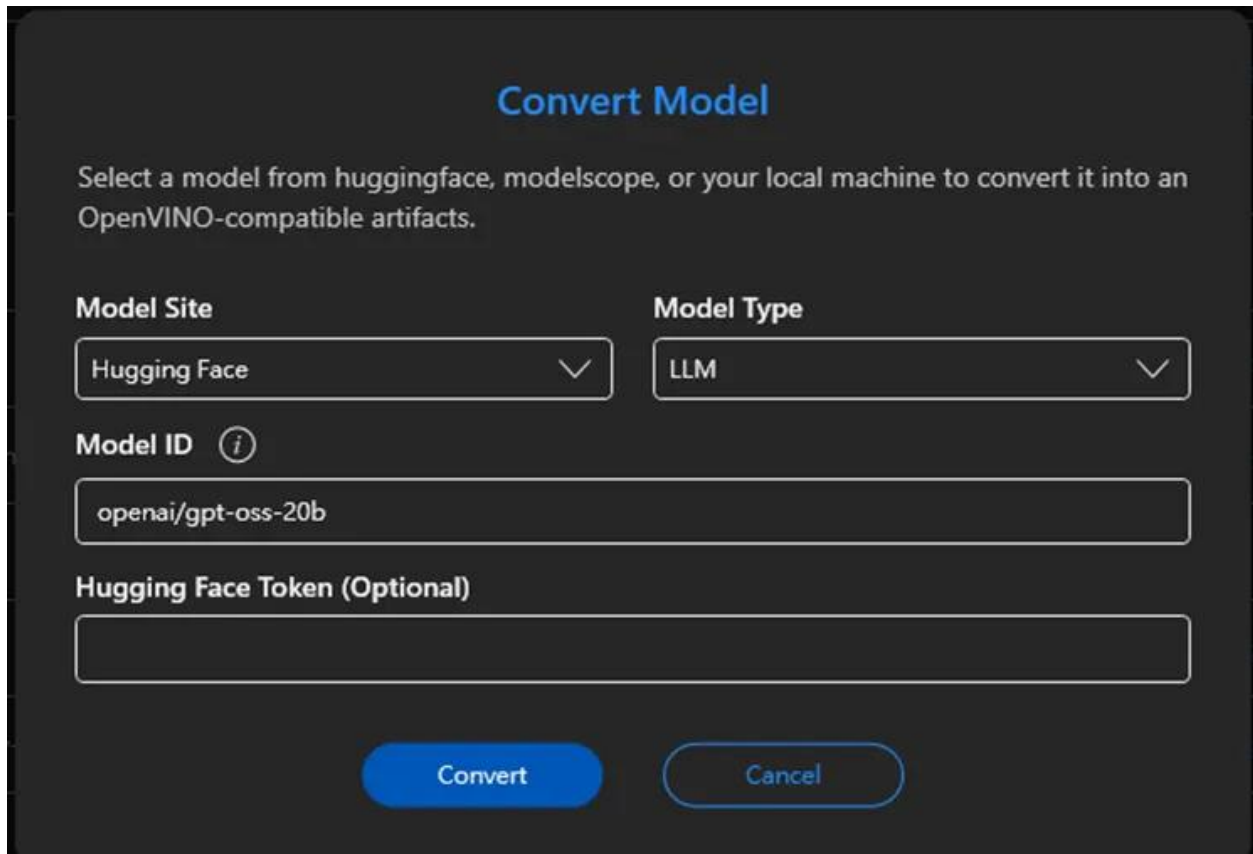
Model Type
LLM
LLM
Embedding
Reranker

Convert Cancel

Start conversion in AI SuperBuild

After you set the **Model Site** and **Model Type**, and provide the correct **Model ID / Path**:

1. Go back to the **Model Upload & Convert** page in AI SuperBuild.
2. Confirm the following settings:
 - **Model Site** is set to **Hugging Face**, **ModelScope**, or **Local Path** as needed.
 - **Model Type** is correctly selected (LLM / Embedding / Reranker).
 - The **Model ID / Path** field contains the copied model ID or local directory path.
3. Click **Convert**.



Convert Model

Select a model from huggingface, modelscope, or your local machine to convert it into an OpenVINO-compatible artifacts.

Model Site **Model Type**

Hugging Face LLM

Model ID ⓘ

openai/gpt-oss-20b

Hugging Face Token (Optional)

Convert Cancel

AI SuperBuild will automatically:

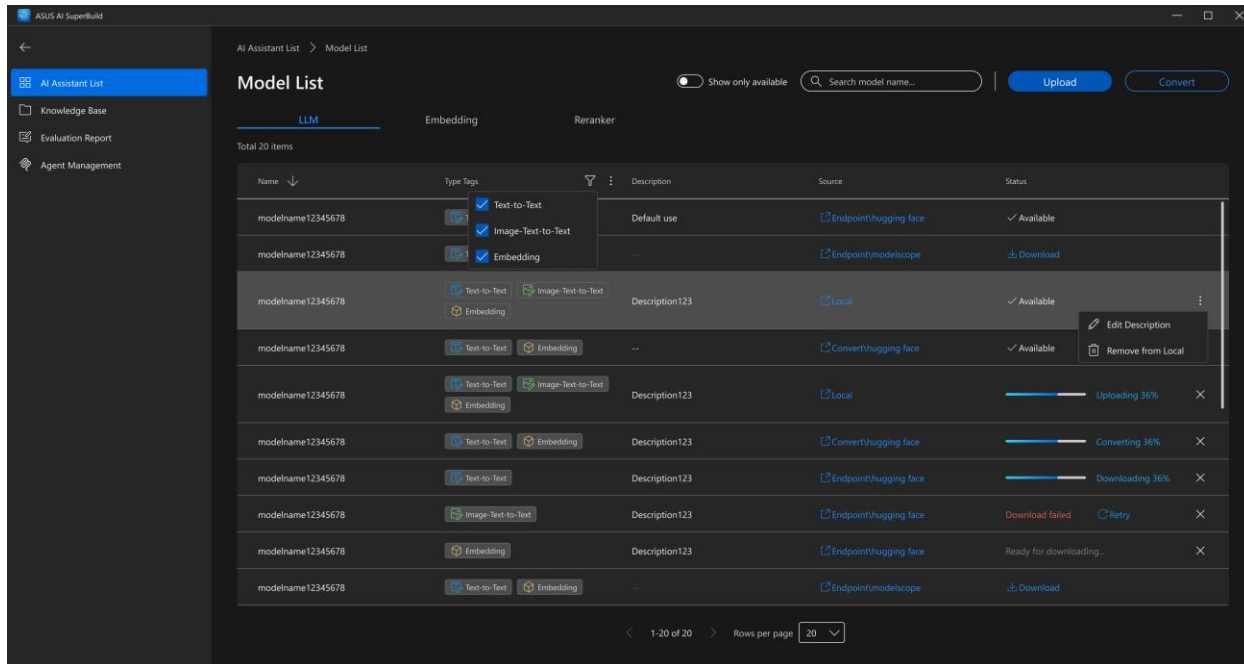
- Download the model from the selected platform (Hugging Face / ModelScope), or load it from the local path.
- Convert it into a format supported by AI SuperBuild.
- Optimize the model for inference (depending on internal configuration).
- Register the converted model into AI SuperBuild so it can be used in projects.

3.2.4 Model Type Tags

To help users quickly grasp the core capabilities of each LLM, the system automatically displays corresponding "Type Tags" right next to each model's name in the Model List. This feature is fully visible by default, allowing users to identify a model's attributes at a glance while browsing the list, without needing any extra clicks or actions.

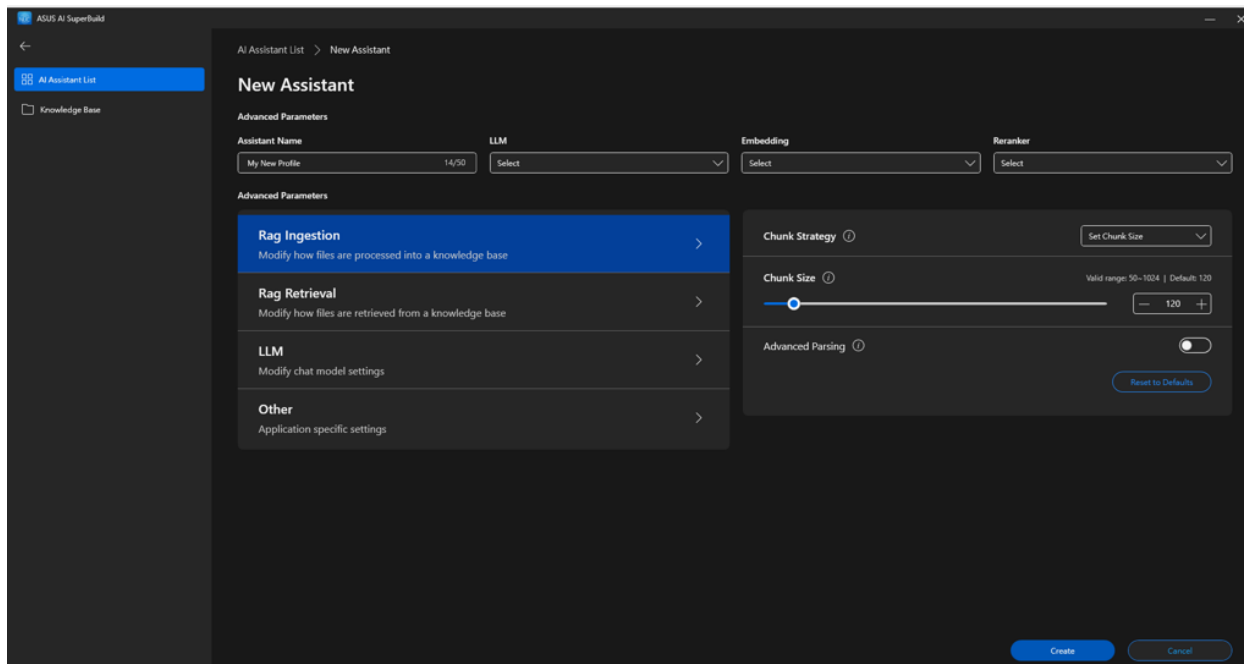
Key Features & Details:

- **Accurate Capability Mapping:** The system ensures that the information displayed on the tags perfectly matches the model's actual supported capabilities, providing users with a reliable reference for selection.
- **Comprehensive Model Types:** The tagging system supports a wide range of mainstream AI capabilities, including but not limited to Text-to-Text, Image-to-Text, and Text-to-Speech.
- **Intuitive Visual Distinction:** To improve search efficiency and user experience, different model types are assigned unique colors or icons. This distinct visual design ensures that users can easily spot and filter the exact type of model they need, even when scrolling through a long list.



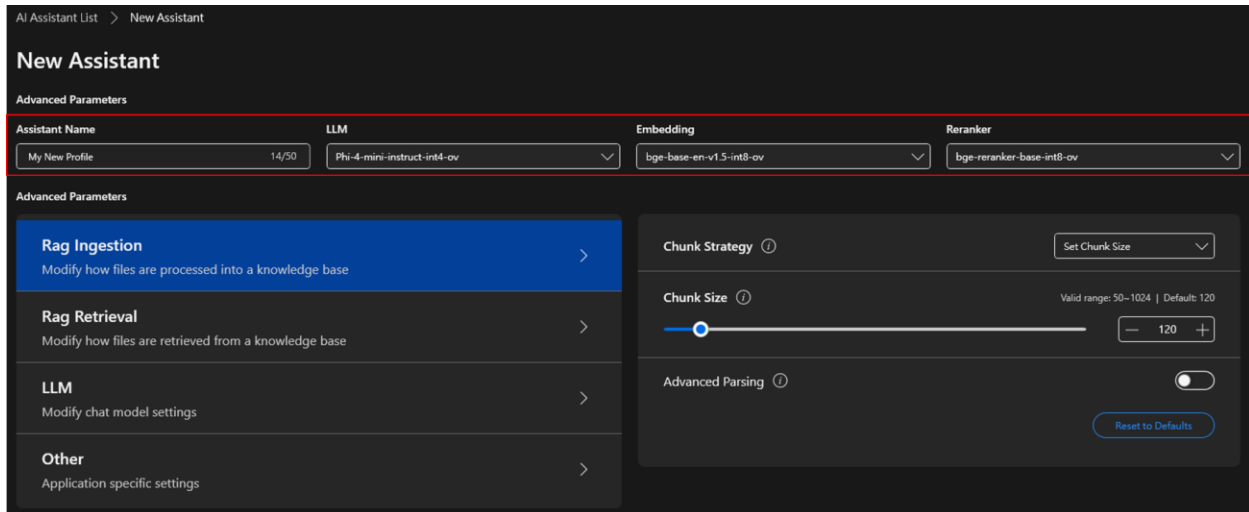
3.3 Creating a New Assistant

When you click "Create New," you are taken to the New Assistant configuration page. This is where you define the core components and behaviors of your new AI assistant, from the base models to the fine-grained parameters that control its performance.



3.3.1 Basic Settings

These are the fundamental components of your AI assistant.



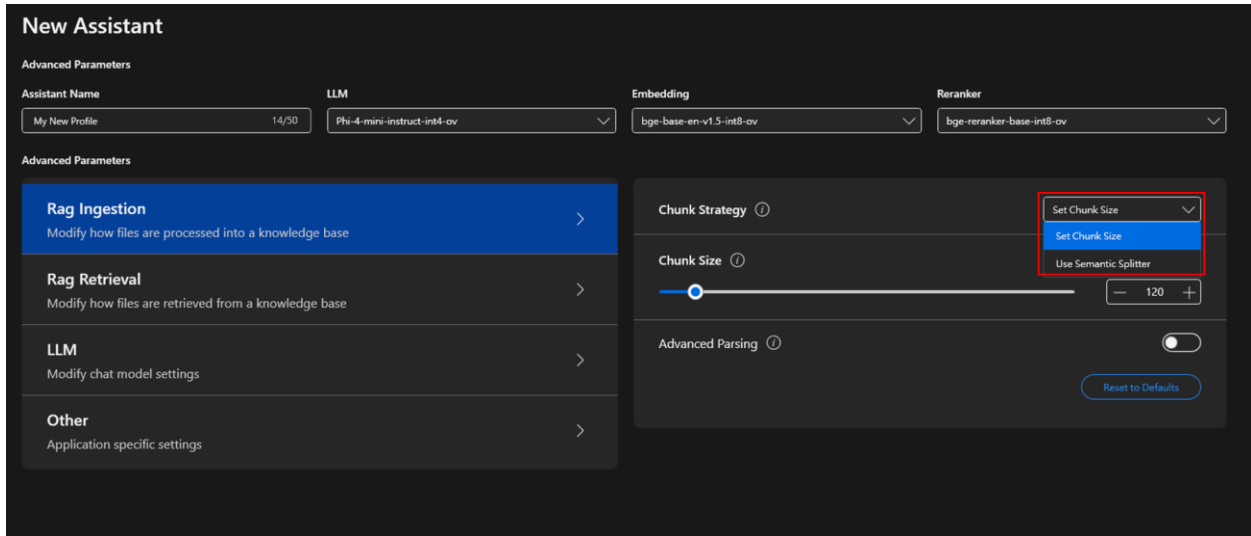
Parameter	Description
Assistant Name	A custom name for your new assistant profile. This name will appear in the Model Selection dropdown on the main chat screen.
LLM (Large Language Model)	The core AI engine that generates human-like text to answer questions, summarize documents, and engage in conversation. From the dropdown, select the base language model that will power this assistant.
Embedding	The model usedis used to convert text into numerical vectors, allowing the AI to understand semantic meaning and relationships. This is crucial for retrieving relevant information from the knowledge base.
Reranker	A secondary model that refines search results by re-evaluating the top documents retrieved by the Embedding model. It improves the accuracy of the context provided to the LLM for generating answers.

3.3.2 Advanced Parameters

This section allows you to fine-tune how the assistant ingests and retrieves information, and how the language model generates its final response.

Rag Ingestion

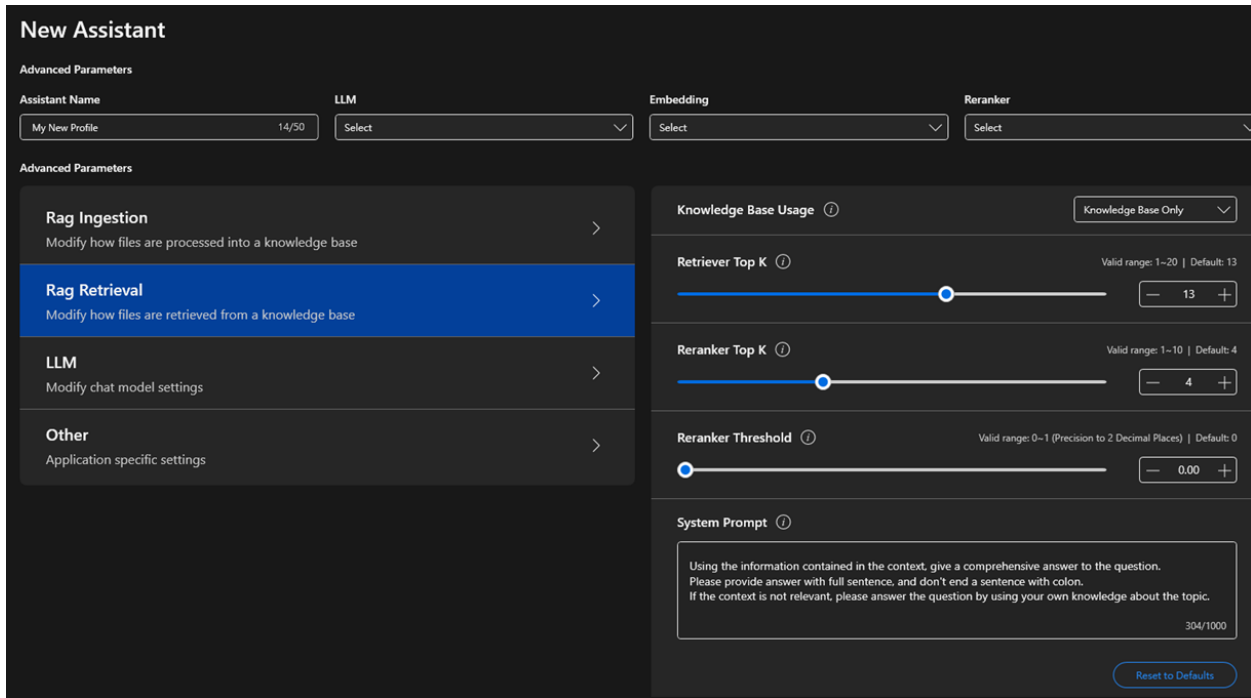
This section controls how your documents are processed and added to the knowledge base.



Parameter	Description
Chunk Strategy	Defines how large documents are split into smaller pieces ("chunks"). This is critical for how the material is indexed and searched. Options include "Set Chunk Size" or "Use Semantic Splitter".
Chunk Size	Sets the size of each text chunk. A smaller size may yield more granular results, while a larger size retains more context. The valid range is 50-1024 (Default: 120).
Advanced Parsing	When enabled, this improves data extraction from complex documents like PDFs by specifically identifying and processing tables, though it may increase upload time.

Rag Retrieval

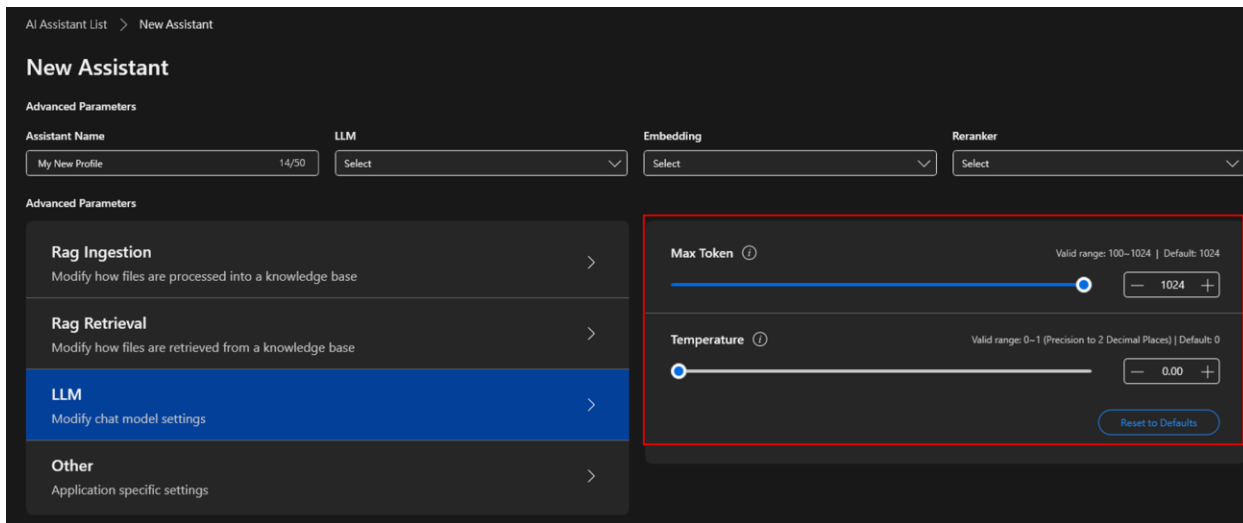
This section defines how the AI searches the knowledge base to find the most relevant information to answer your queries.



Parameter	Description
Retriever Top K	Controls how many of the most relevant documents ("chunks") are initially fetched from the knowledge base. A higher number provides more context but may increase processing time.
Reranker Top K	Determines how many of the initially retrieved documents are passed to the reranker for a more detailed relevance scoring. The reranker selects the best documents from this smaller subset.
Reranker Threshold	Sets a minimum relevance score for a reranked document to be considered valid context. Documents scoring below this are discarded.
System Prompt	A set of instructions that guides the LLM on its behavior, such as how to use the retrieved context, what tone to adopt, or what to do if the context is not relevant.

LLM

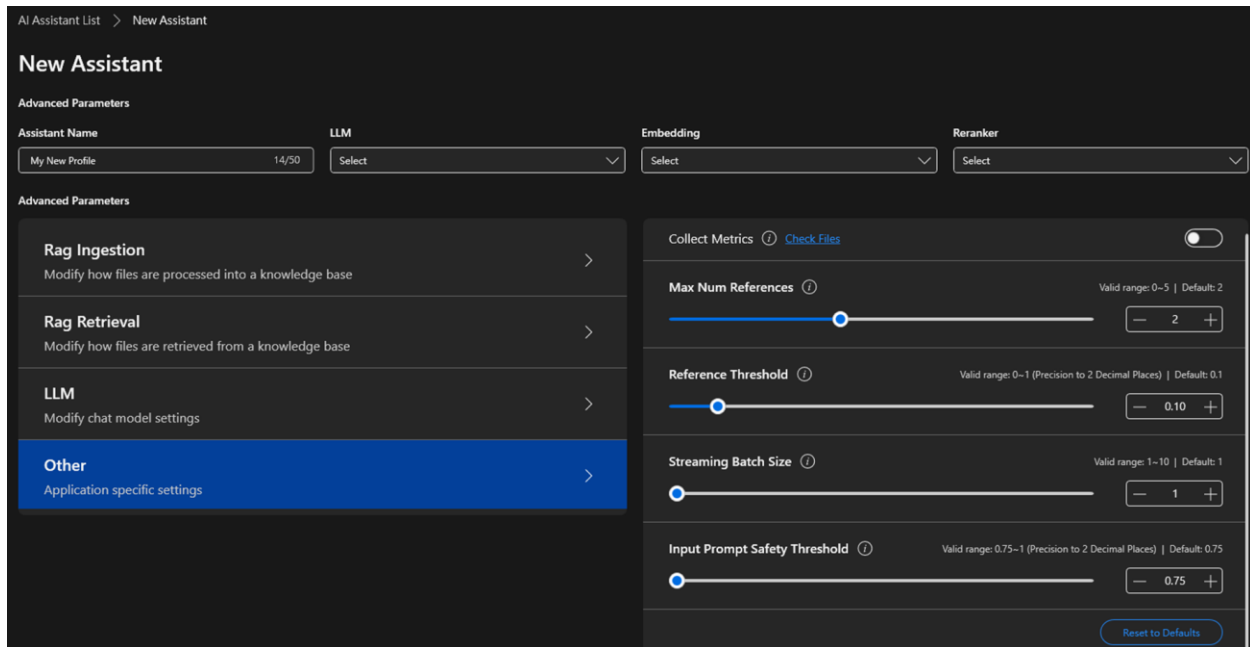
Here you can modify the chat model's settings to control the style and length of the AI's responses.



Parameter	Description
Max Token	Sets the maximum number of tokens (words or parts of words) that the model can generate in a single response, effectively controlling the answer's maximum length.
Temperature	Controls the randomness of the model's output. A lower value (e.g., 0.1) makes responses more deterministic and factual, while a higher value (e.g., 0.9) encourages more creativity.

Other

This section contains various application-specific settings for advanced control over the assistant's behavior and performance monitoring.



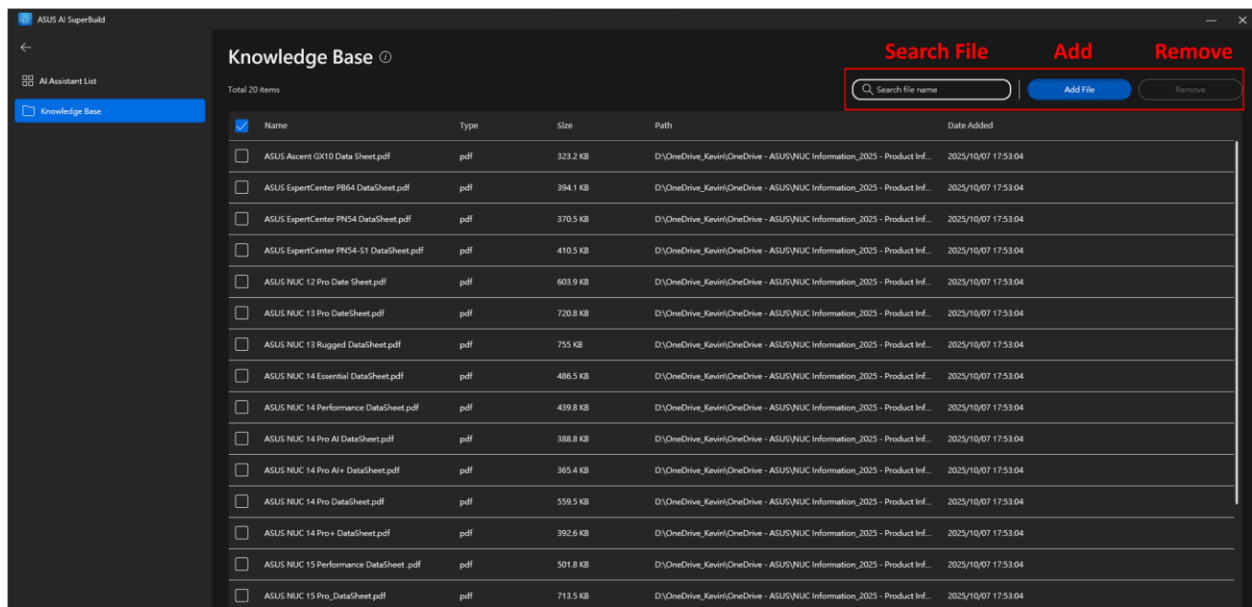
Parameter	Description
Collect Metrics	Enables or disables the collection of performance and usage data for analysis.
Max Num References	Sets the maximum number of source documents from the knowledge base that can be cited in a single response.
Reference Threshold	Defines the minimum relevance score a document must have to be considered as a citable reference in an answer.
Streaming Batch Size	Controls the number of tokens processed in each batch when generating a response in streaming mode.
Input Prompt Safety Threshold	Sets the confidence level for the content safety filter applied to user inputs. Prompts deemed unsafe will be blocked.

3.4 Editing an Existing Assistant

In addition to creating new assistants, you can easily modify existing configurations. On the **AI Assistant List** page, hover over the assistant you wish to modify, click the "..." icon that appears on the right, and select **Edit**. This will open the same detailed parameter page as when creating a new assistant, allowing you to fine-tune all basic and advanced settings for that specific assistant.

3.5 Knowledge Base Management

The Knowledge Base is where you manage the source files for Retrieval-Augmented Generation (RAG). By uploading documents (PDF, DOCX, TXT, PPTX.), you provide the AI assistant with a specific set of information to draw from, ensuring more accurate and context-aware responses grounded in your data.



Button/UI Element	Description
Add File	Opens a file browser to upload one or more supported documents to the knowledge base.
Remove	Deletes the selected file(s) from the knowledge base.
Search for file name...	A search bar to quickly find specific files in the knowledge base by name.

3.6 Evaluation and Report

Determining the optimal combination of LLMs and Knowledge Base files can be challenging. The Evaluation feature streamlines this process by benchmarking your system's performance against a standard dataset.

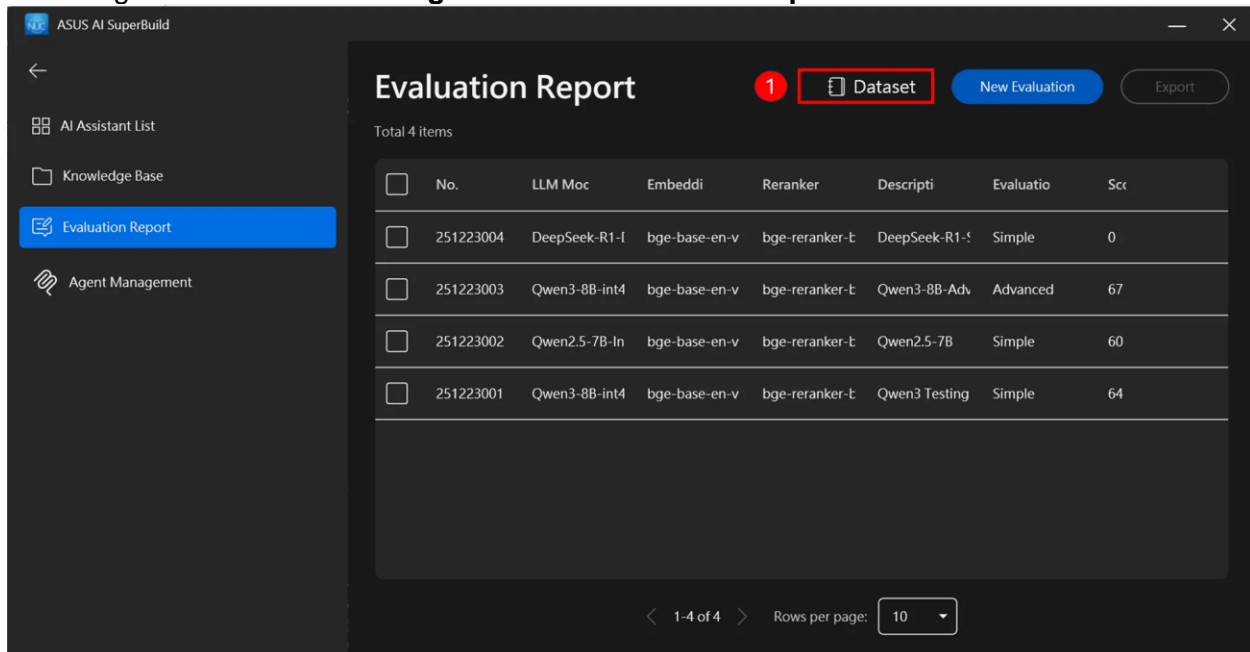
Two evaluation modes are available:

- **Simple Mode:** Quickly calculates an accuracy score based on uploaded question-answer pairs to benchmark the current LLM and Knowledge Base.
- **Advanced Mode:** Leverages cloud-based AI to generate comprehensive reports with actionable improvement suggestions.

3.6.1 Preparing the Evaluation Dataset

To ensure data privacy and relevance, you must provide a custom evaluation dataset containing domain-specific questions and ground-truth answers. This dataset serves as the benchmark for performance evaluation.

1. Navigate to **Assistant Configuration >> Evaluation Report >> Dataset**

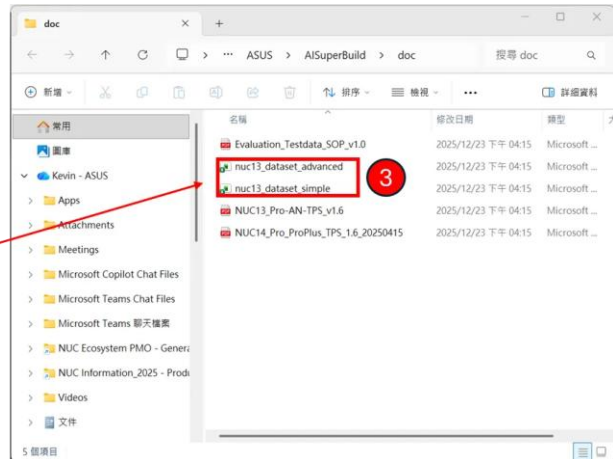
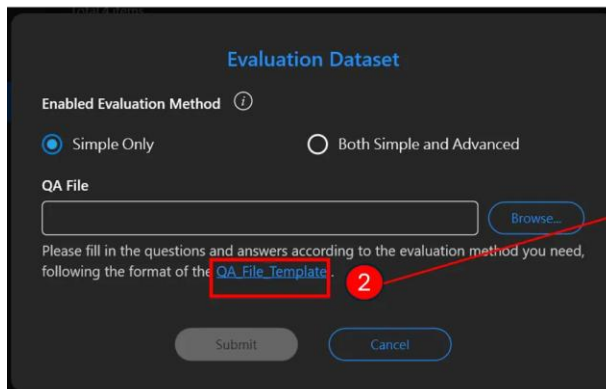


The screenshot shows the 'Evaluation Report' interface in the ASUS AI SuperBuild application. The left sidebar contains navigation options: AI Assistant List, Knowledge Base, Evaluation Report (highlighted), and Agent Management. The main content area is titled 'Evaluation Report' and shows a table with 4 items. A red box highlights the 'Dataset' button in the top right corner. The table has columns for No., LLM Model, Embedding, Reranker, Description, Evaluation Mode, and Score.

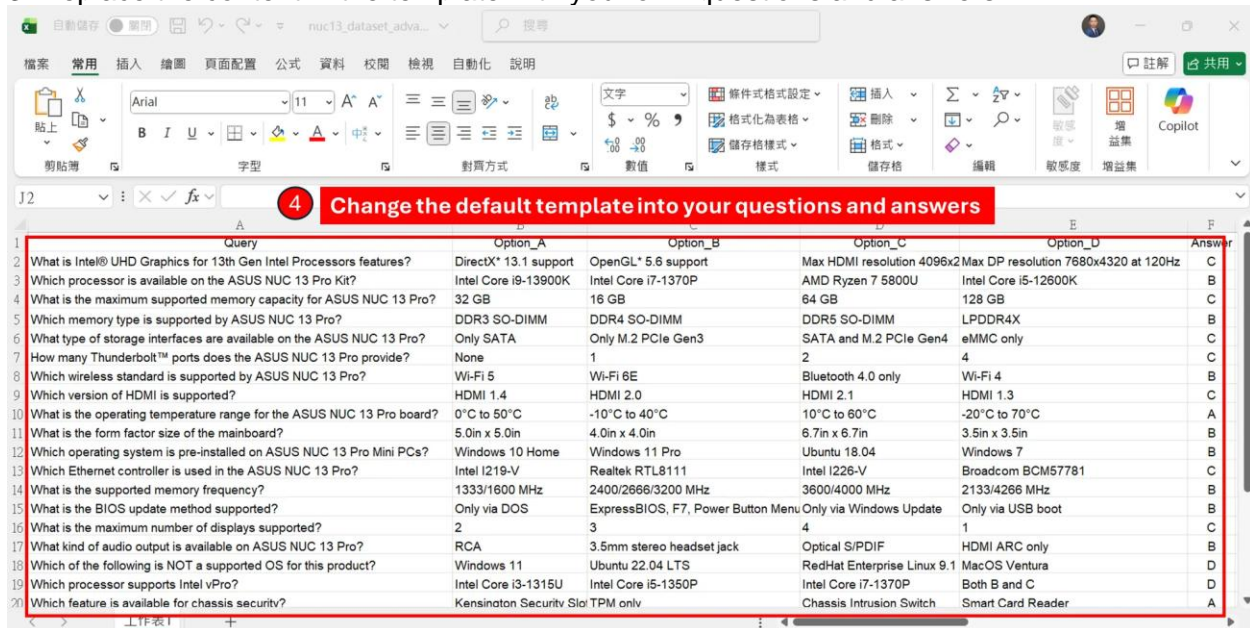
No.	LLM Model	Embedding	Reranker	Description	Evaluation Mode	Score
251223004	DeepSeek-R1-T	bge-base-en-v	bge-reranker-t	DeepSeek-R1-T	Simple	0
251223003	Qwen3-8B-int4	bge-base-en-v	bge-reranker-t	Qwen3-8B-Adv	Advanced	67
251223002	Qwen2.5-7B-In	bge-base-en-v	bge-reranker-t	Qwen2.5-7B	Simple	60
251223001	Qwen3-8B-int4	bge-base-en-v	bge-reranker-t	Qwen3 Testing	Simple	64

At the bottom of the table, there is a pagination control showing '1-4 of 4' and a 'Rows per page' dropdown set to '10'.

2. Click **QA_File_Template** to download the example file.



3. Replace the content in the template with your own questions and answers.



4. Click "Change" to upload your updated dataset file, then click "Submit".

Evaluation Dataset

Enabled Evaluation Method (i)

Simple Only Both Simple and Advanced

QA File **5** Upload the dataset

C:\Program Files\ASUS\AISuperBuild\doc\nuc13_dataset_advanced.xlsx Change

Please fill in the questions and answers according to the evaluation method you need, following the format of the [QA File Template](#).

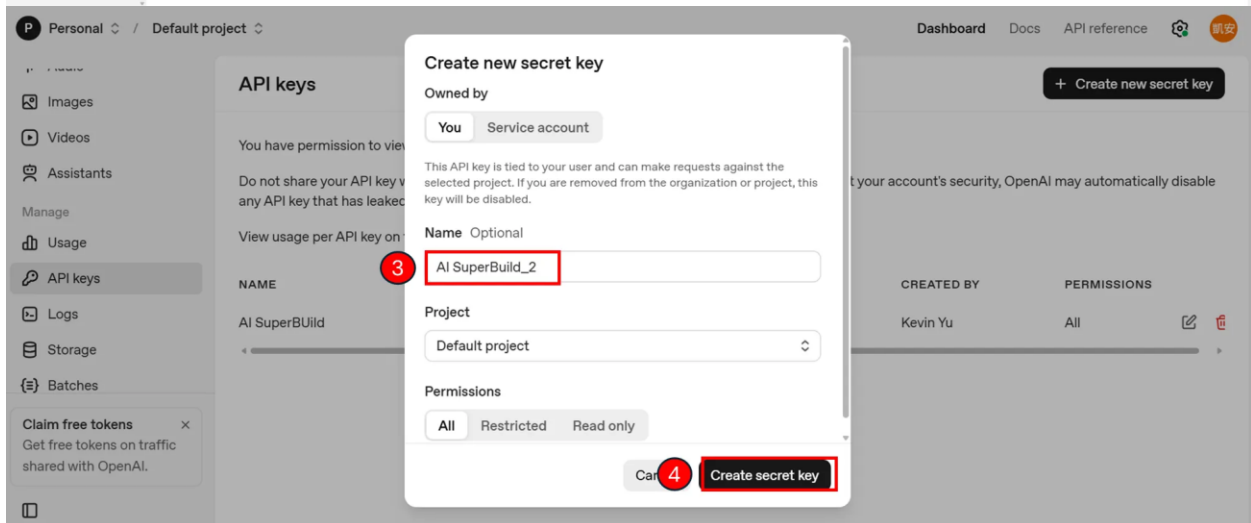
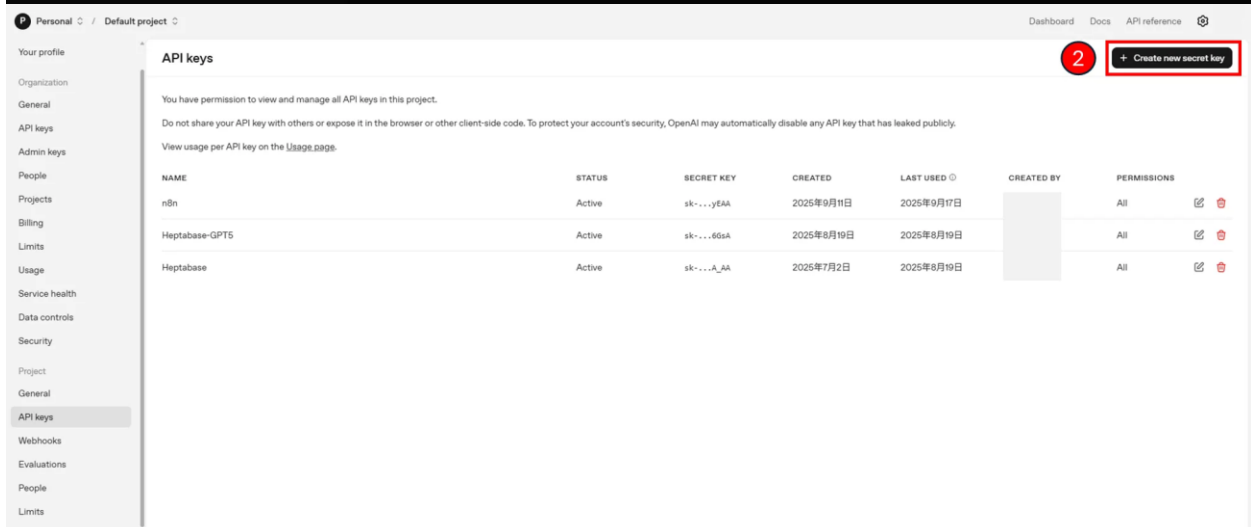
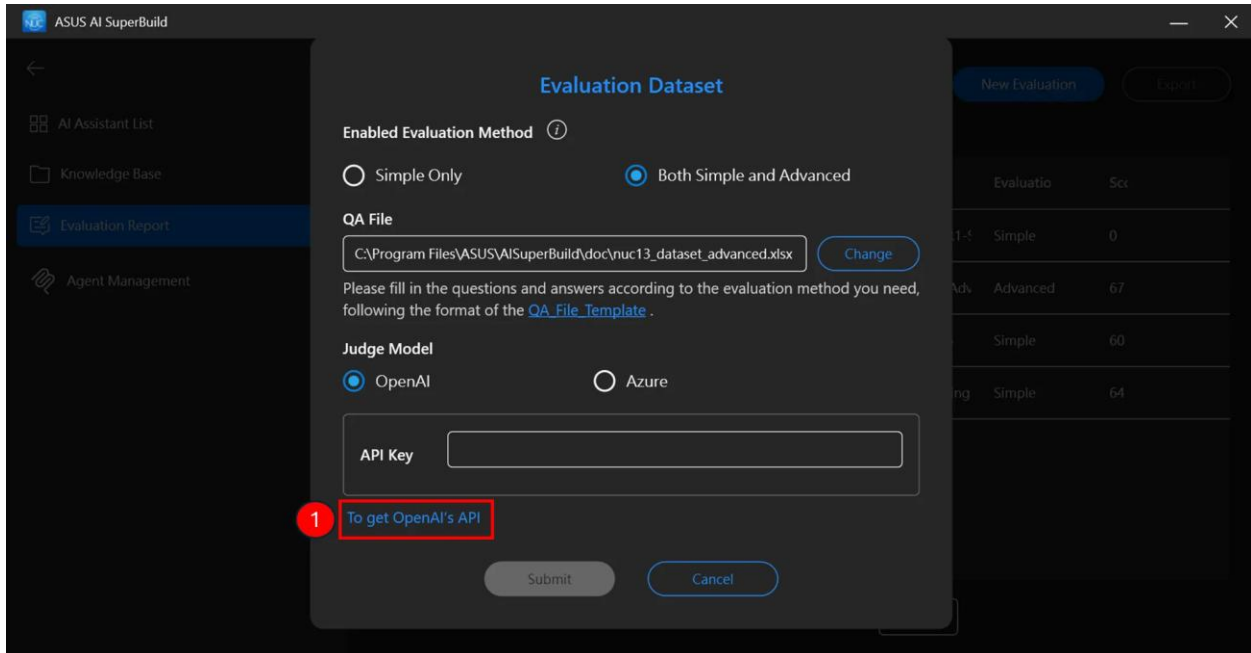
Submit Cancel

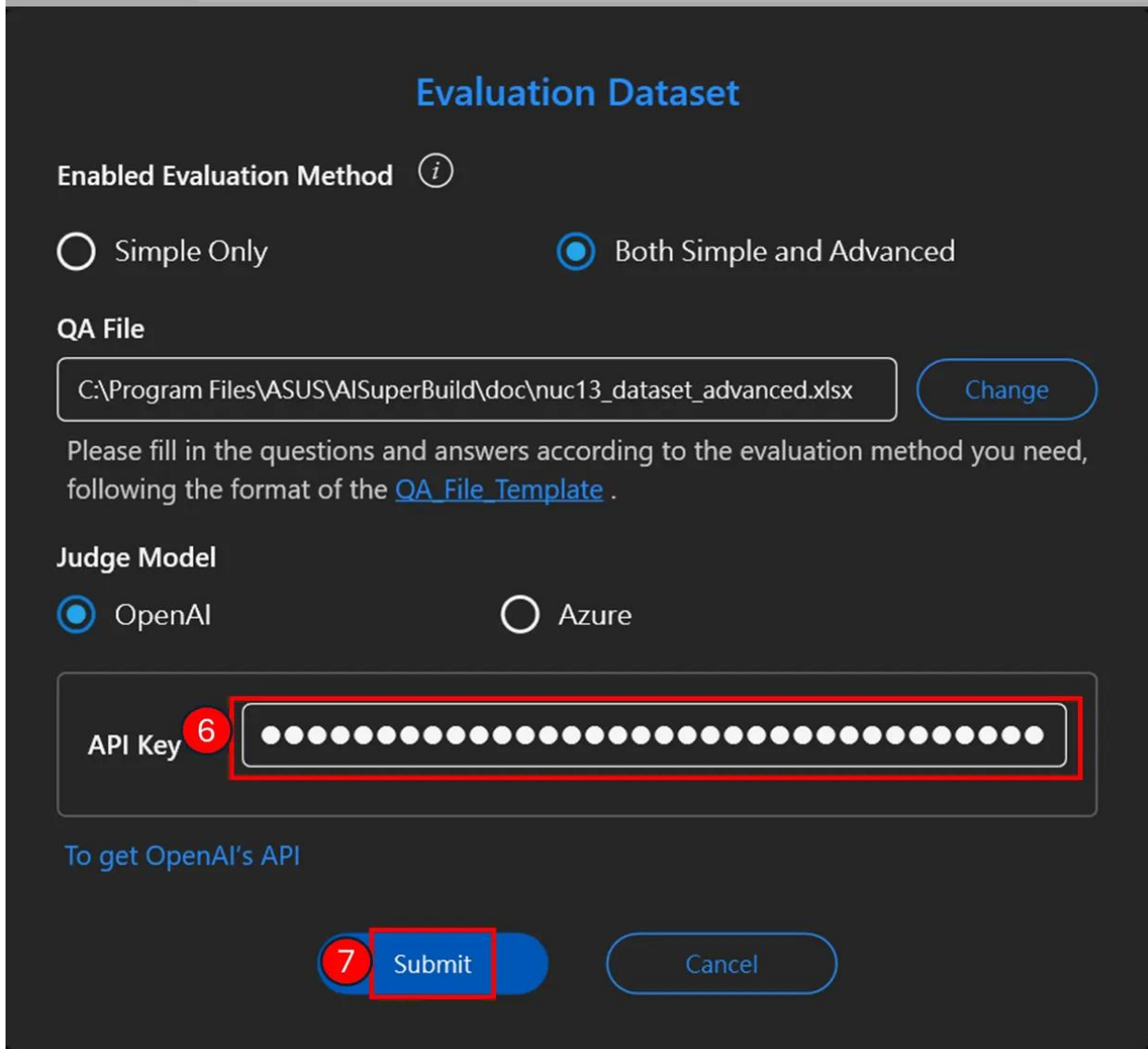
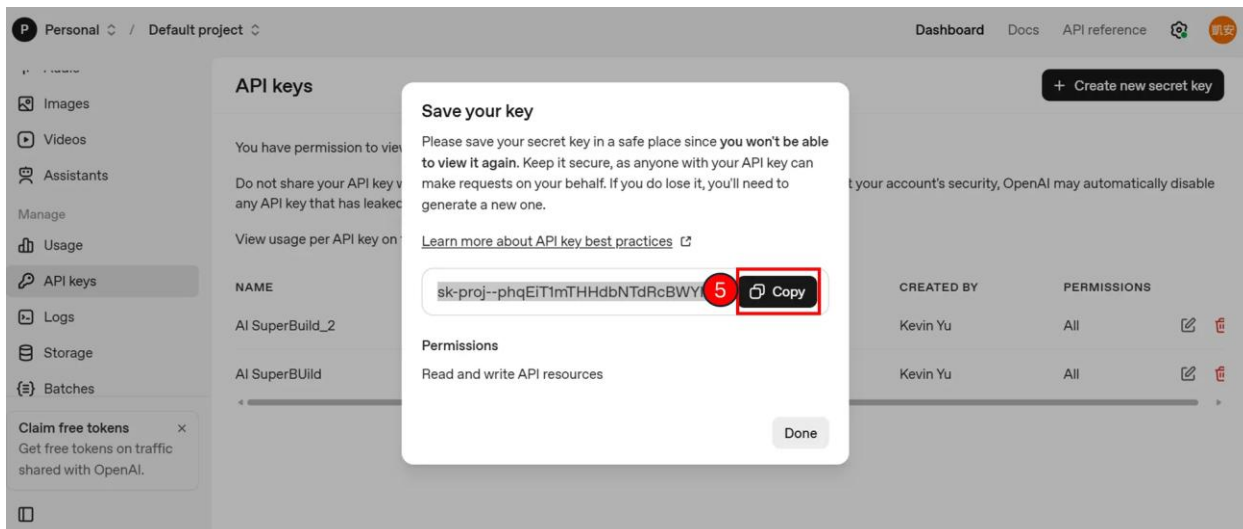
Evaluation Configuration

- For basic benchmarking, select **Simple Only**.
- For detailed analysis, select **Both Simple and Advanced**. This requires a cloud-based LLM API key (OpenAI or Azure OpenAI).

OpenAI API Configuration

1. Click the **To get OpenAI's API** link to obtain your API key.
2. Enter the key into the **API Key** field in AI SuperBuild.





Azure OpenAI

To use Azure OpenAI with AI SuperBuild, you must provide both your **API key** and **endpoint**. You can obtain this information from the Azure OpenAI portal (ai.azure.com). After that, create and deploy the required model (currently, only **gpt-4.1-mini** is supported for evaluation). Follow the steps below in the Azure AI portal:

1. Go to the **Develop** section to find your **Key** and **Endpoint**.

[Find Key and Endpoint in Develop]

The screenshot shows the Microsoft Azure portal interface. At the top, there's a search bar and a 'Copilot' button. Below that, the 'Azure OpenAI' resource overview is displayed. The 'Essentials' section shows details like Resource group, Status (Active), Location (East US), and Subscription ID. On the right, it lists API Kind (OpenAI), Pricing tier (Standard), Endpoints (with a link to view endpoints), and Manage keys (with a link to manage keys). The 'Keys and endpoint' section is active, showing a 'Show Keys' button and two keys (KEY 1 and KEY 2) with their respective endpoints. The Location/Region is set to 'eastus'.

2. Go to **Get started** and enter the **Explore** experience to deploy a model.
[Enter Explore and deploy in Get started]

The screenshot displays the Microsoft Azure portal interface. At the top, the header includes the Microsoft Azure logo, a search bar, and a Copilot icon. Below the header, there are navigation tabs for 'Home', 'Azure OpenAI', and several AI-related articles. The main content area is divided into a left-hand navigation pane and a central content pane. The navigation pane lists various services and management tools, with 'Overview' selected. The central content pane features a blue banner with a message about upgrading to Foundry for access to third-party models. Below this, the 'Essentials' section provides key resource details: Resource group (mcsd), Status (Active), Location (East US), Subscription (mcsd), and Subscription ID. A 'Tags' section is also visible. The main content area is titled 'Build your own secure copilot and generative AI applications with Azure OpenAI Service' and includes a 'Get Started' button.

Microsoft Azure

Search resources, services, and docs (0+)

Copilot

Home

Azure OpenAI

How to use this Azure AI services with Python

Best practices for managing AI resources

Which npm package should I install to use this AI resource?

Search

Go to Foundry portal

Delete

Want to try the latest industry models and Agents?

Azure OpenAI is limited to OpenAI models. Upgrading your resource type to Foundry gives you access to Open AI plus third-party models like Meta, Mistral, vAI and more. Build enterprise grade AI agents using your data and tools.

Get Started

Essentials

Resource group (mcsd)

Status : Active

Location : East US

Subscription (mcsd) :

Subscription ID :

Tags (add) : Add tags

API Kind : OpenAI

Pricing tier : Standard

Endpoints : [Click here to view endpoints](#)

Manage keys : [Click here to manage keys](#)

Get Started

Develop

Monitor

Build your own secure copilot and generative AI applications with Azure OpenAI Service

Deploy an Azure OpenAI model and start making API calls. Connect your own data, call functions, and improve workflow with Azure OpenAI language, image and speech models. You can access the service through REST APIs, Python SDK, or our web-based interface in the Foundry portal.

Learn More

Explore and deploy

Explore and deploy the generative AI models, craft unique prompts for your use cases, and fine-tune select models.

Explore Foundry portal

3. In the **Chat** section, select **Create a deployment**.
[Tap Create a deployment in Chat]

The screenshot shows the Microsoft Foundry Chat playground interface. The left sidebar contains navigation options: Home, Get started, Model catalog, Playgrounds, Chat (selected), Assistants (PREVIEW), Video (PREVIEW), Audio (PREVIEW), Images, Tools (Fine-tuning, Azure OpenAI Evaluation (PREVIEW), Stored completions (PREVIEW), Batch jobs, Monitoring), and Shared resources (Deployments, Quota, Guardrails + Controls, Risks + alerts (PREVIEW), Data files, Assistant vector stores (PREVIEW)). The main content area is titled 'Chat playground' and includes a toolbar with 'View code', 'Deploy', 'Import', 'Export', 'Prompt samples', and 'Filter'. The 'Setup' section is active, showing 'Deployment *' with a dropdown menu open. The dropdown menu has three options: 'Create new deployment', 'From base models' (which is selected), and 'From fine-tuned models'. Below the dropdown, there is a large blue folder icon with a white plus sign. The text reads: 'Deployment needed. In order to modify and interact with the Playground, you first need to deploy a base model to your project. Don't have a deployment? + Create a deployment'.

4. Choose **gpt-4.1-mini** as the model and confirm.
[Select gpt-4.1-mini and tap Confirm]

Select a model

Choose a model to create a new deployment. For flows and other resources, create a deployment from their respective list. [Go to model catalog.](#)

Models 27 Inference tasks: Chat completion Show description

- gpt-5**
Chat completion, Responses
- gpt-4.1**
Chat completion, Responses
- gpt-4.1-mini**
Chat completion, Responses
- gpt-5-codex**
Chat completion, Responses
- o3**
Responses, Chat completion

< Prev
Next >

gpt-4.1-mini

Task: Chat completion

Task: Responses

Direct from Azure models

Direct from Azure models are a select portfolio curated for their market-differentiated capabilities:

- Secure and managed by Microsoft: Purchase and manage models directly through Azure with a single license, consistent support, and no third-party dependencies, backed by Azure's enterprise-grade infrastructure.
- Streamlined operations: Benefit from unified billing, governance, and seamless PTU portability across models hosted on Azure - all as part of one Azure AI Foundry platform.
- Future-ready flexibility: Access the latest models as they become available, and easily test, deploy, or switch between them within Azure AI Foundry; reducing integration effort.
- Cost control and optimization: Scale on demand with pay-as-you-go flexibility or reserve PTUs for predictable performance and savings.

[Learn more about Direct from Azure models.](#)

Confirm

Cancel

5. Configure the deployment:

- **Deployment type** → *Standard*
- **Model version upgrade policy** → *Opt out of automatic model version upgrades*

Then select **Deploy**.

[Select Deployment type → Standard, Model version upgrade policy → Opt out of automatic model version upgrades, Tap Deploy]

Deploy gpt-4.1-mini

Deployment name *



gpt-4.1-mini

Deployment type

Standard

Standard: Pay per API call with lower rate limits. Adheres to Azure data residency promises. Best for intermittent workloads with low to medium volume. Learn more about [Standard deployments](#).

Deployment details

Collapse

Model version upgrade policy

Opt out of automatic model version upgrades

Model version

2025-04-14 (Default)

AI resource

200K tokens per minute quota available for your deployment

Tokens per Minute Rate Limit

100K

Corresponding requests per minute (RPM) = 100

Content filter

DefaultV2

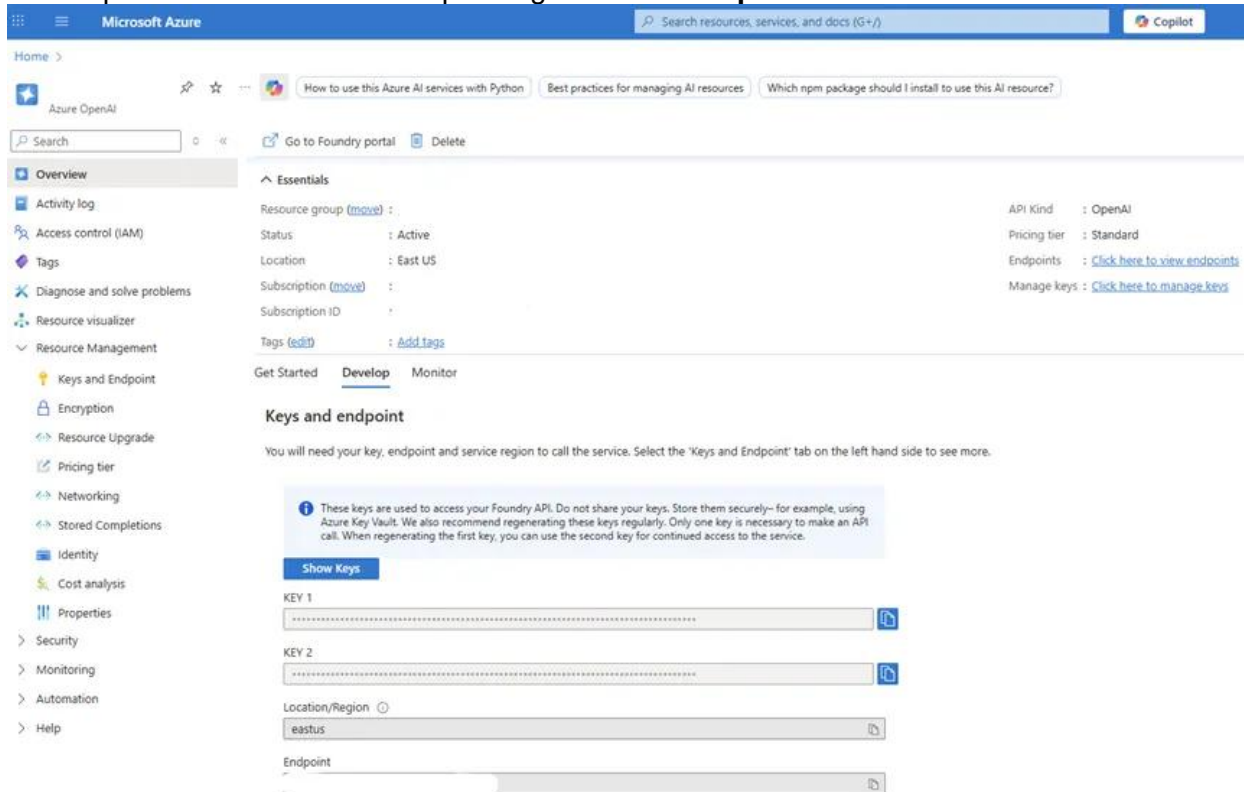
Enable dynamic quota

Enabled

Deploy

Cancel

6. After the deployment is created, copy the **Key** and **Endpoint** from the Azure portal and paste them into the corresponding fields in **AI SuperBuild**.



Evaluation Dataset

Enabled Evaluation Method **Model ID** (i)

Simple Only Both Simple and Advanced

QA File

Browse...

Please fill in the questions and answers according to the evaluation method you need, following the format of the [QA File Template](#).

Judge Model

OpenAI Azure

Endpoint

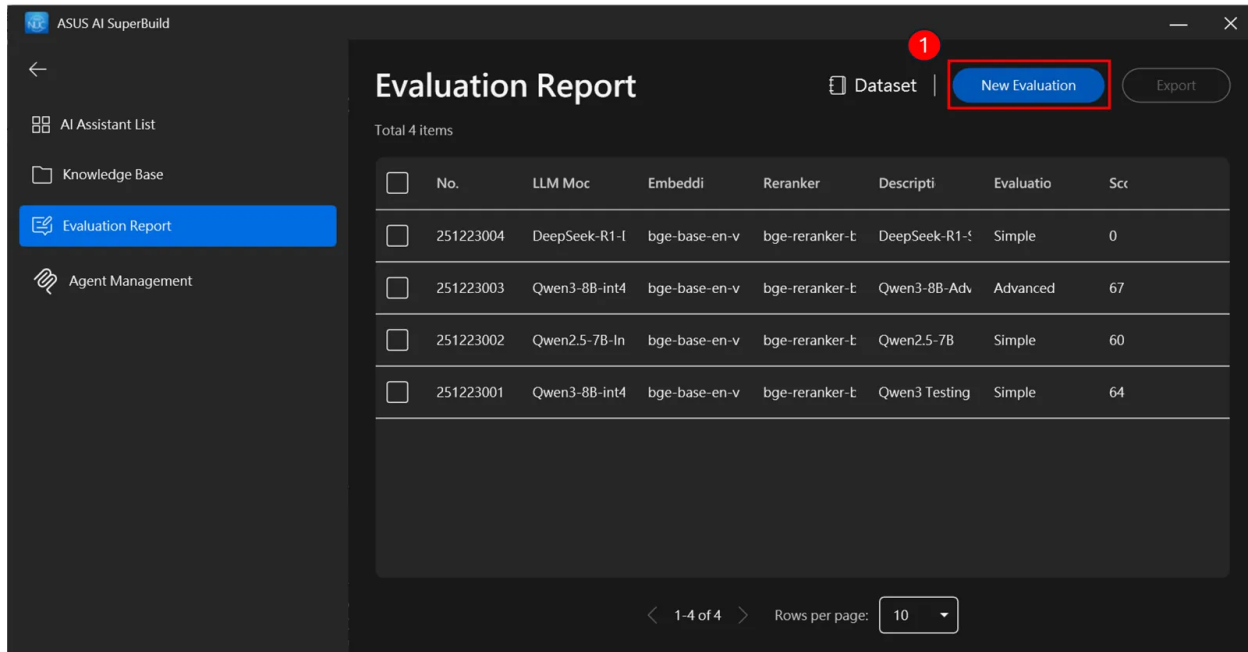
API Key

Learn how to get Azure API information in our [User Guide](#).

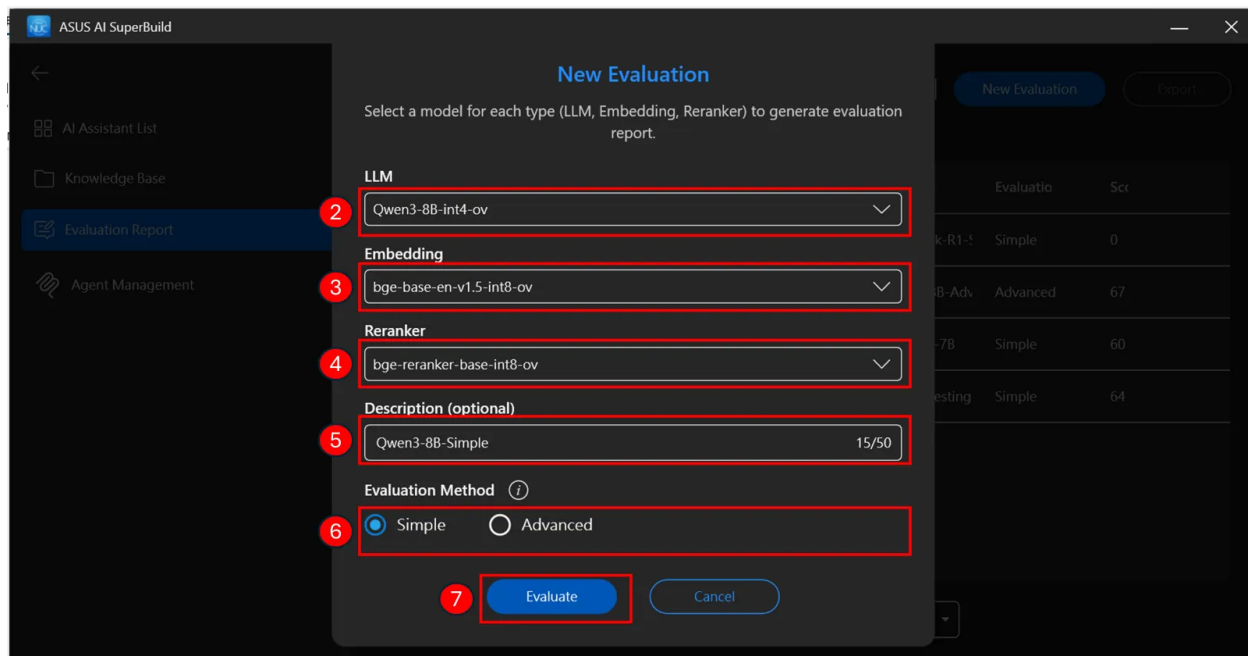
Submit Cancel

3.6.2 Starting the Evaluation

After the dataset is configured, you can evaluate whether the **LLM**, **Embedding**, and **Reranker** models—together with the knowledge base—are sufficient for the target customer scenario. Navigate to "**New Evaluation**".

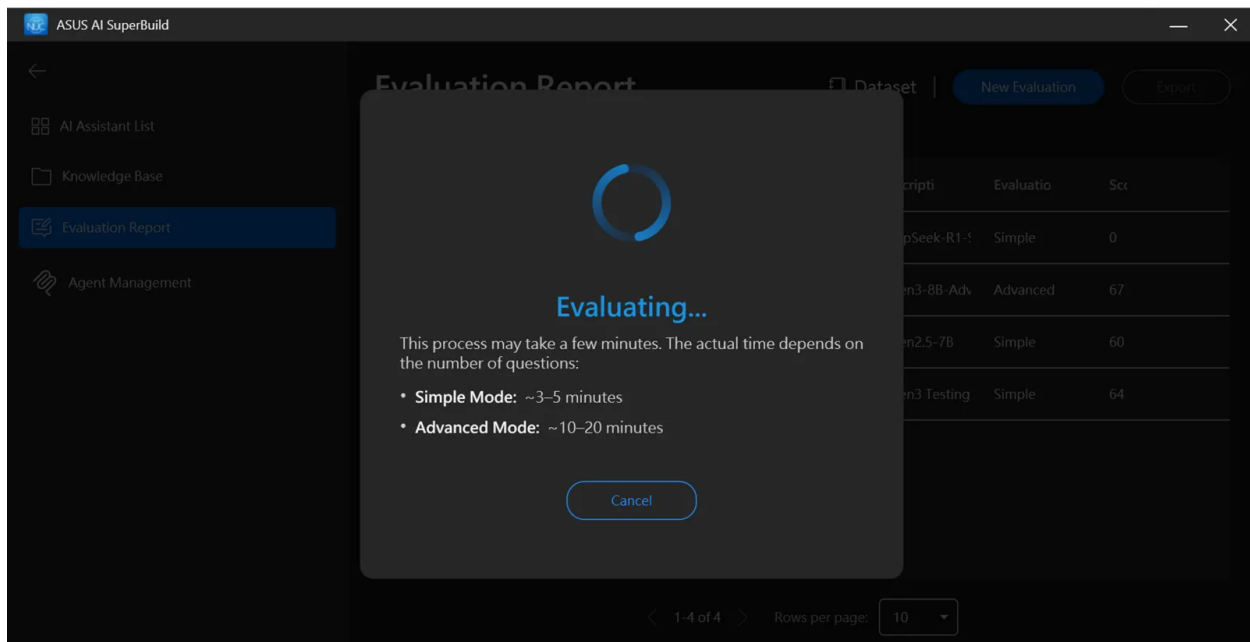


Select the **LLM**, **Embedding**, and **Reranker** models you want to use.

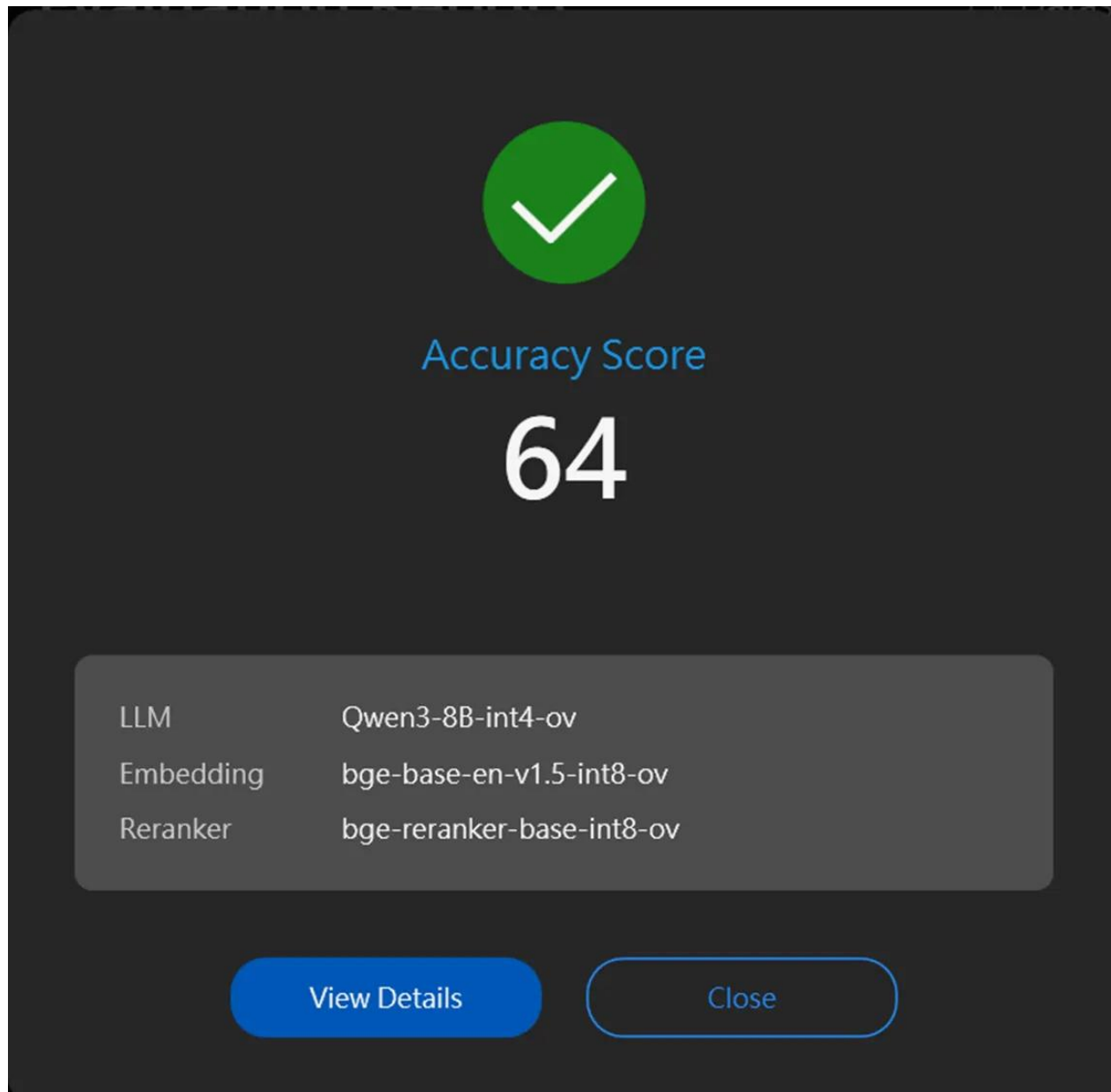


Based on the selected evaluation mode, the system begins processing the results.

- **Simple mode** typically takes around **3–5 minutes**.
- **Advanced mode** typically takes around **10–20 minutes**, depending on the number of questions and the selected models.

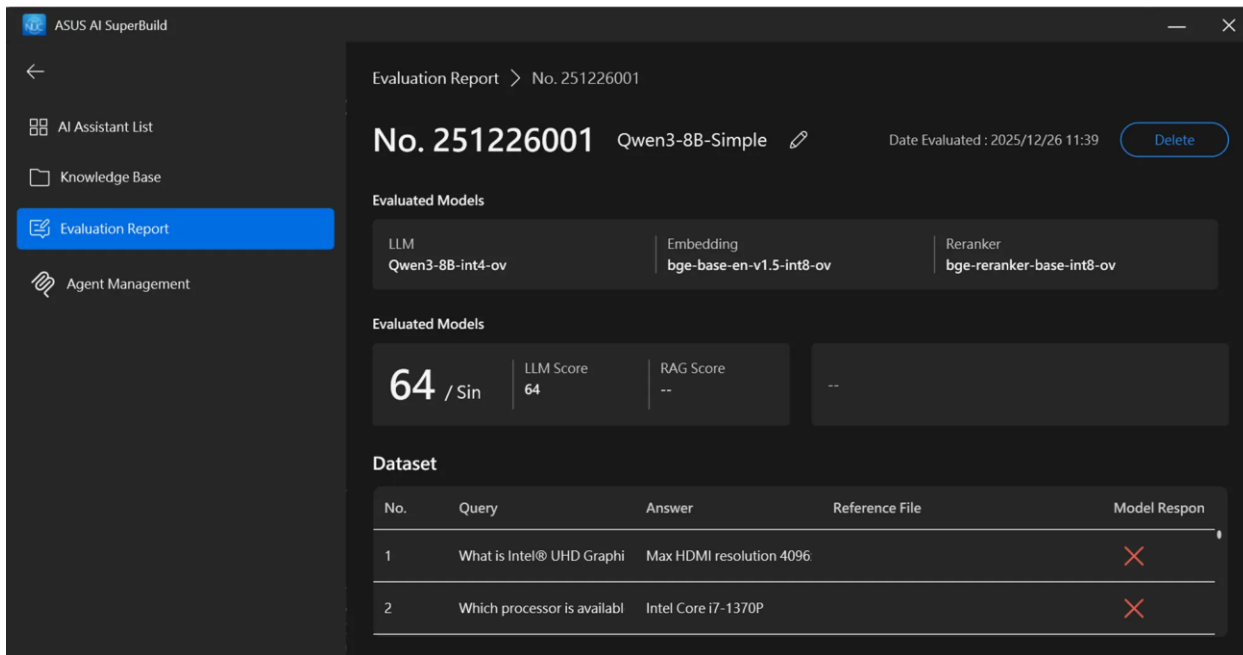


After the evaluation is complete, AI SuperBuild returns an accuracy score that summarizes the overall performance.



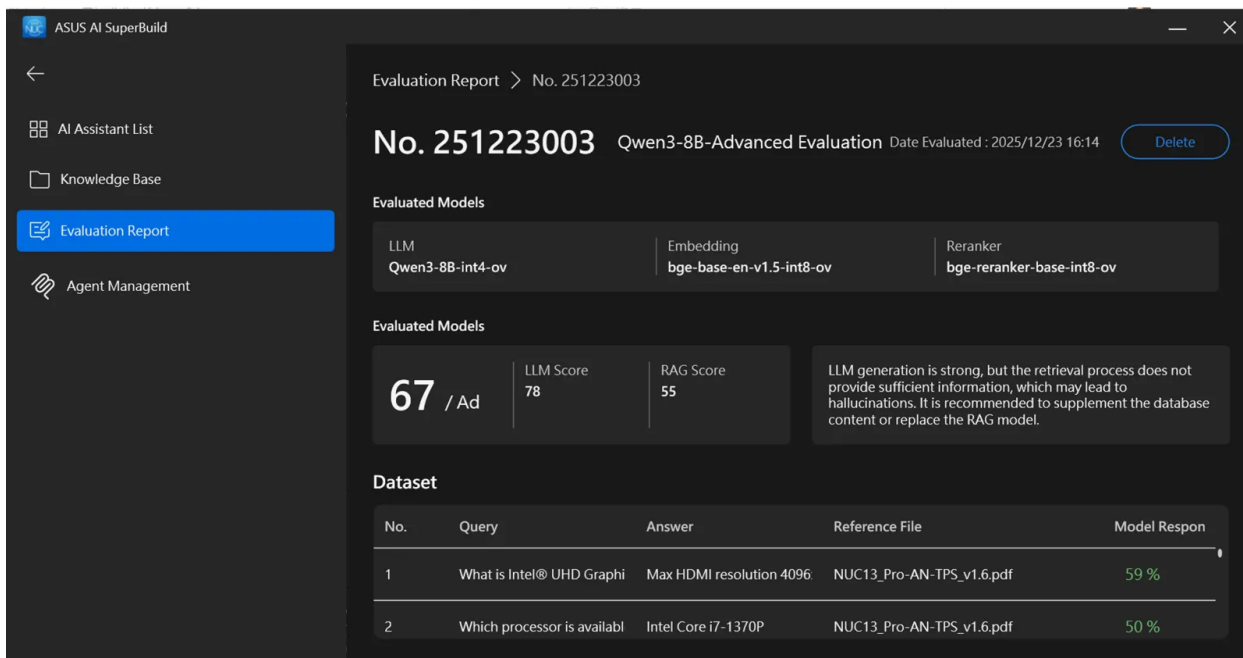
Simple mode results show:

- The overall score
- Whether each question is answered correctly or not



Advanced mode results show:

- The overall accuracy
- Detailed improvement suggestions, such as:
 - Switching to a more powerful model
 - Adding more or better-curated knowledge base files

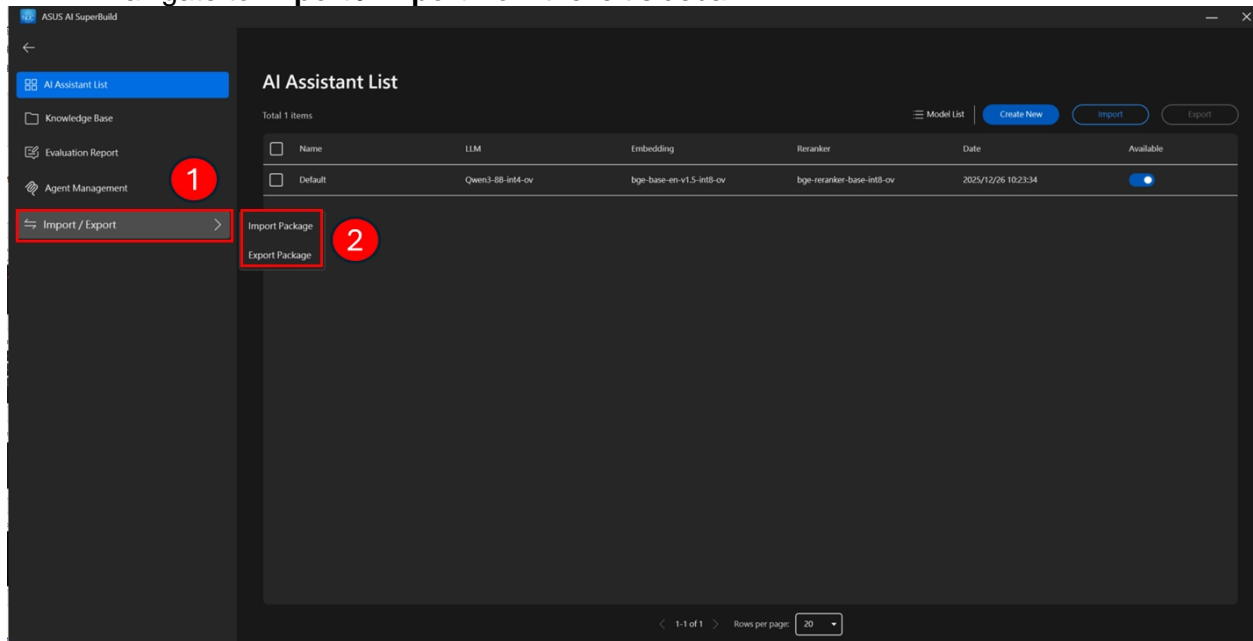


3.7 One-Key Import/Export

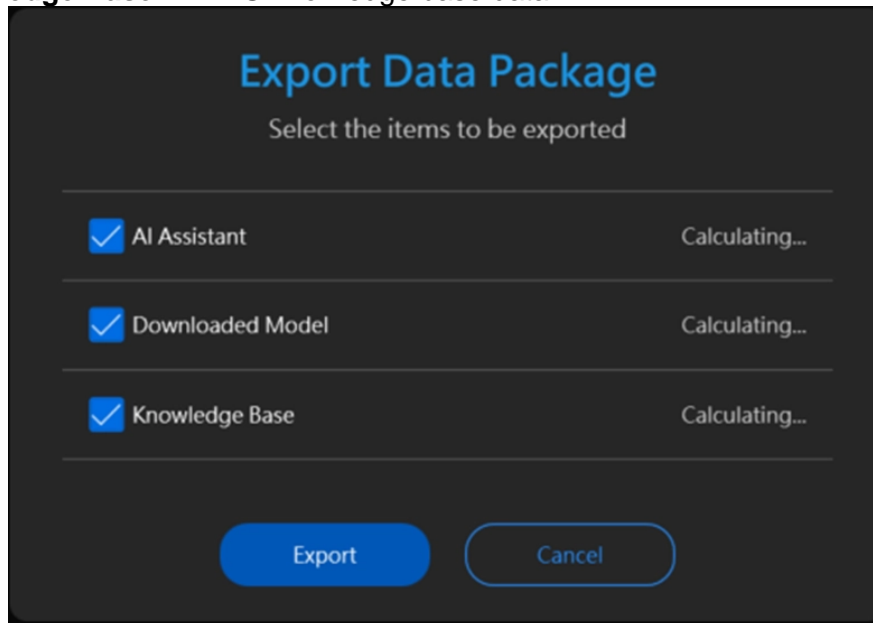
This feature allows users to back up and restore all configuration files with a single click, making it easy to migrate settings across devices or recover from changes.

3.7.1 Export Configuration

1. Navigate to **Import / Export** from the left sidebar.

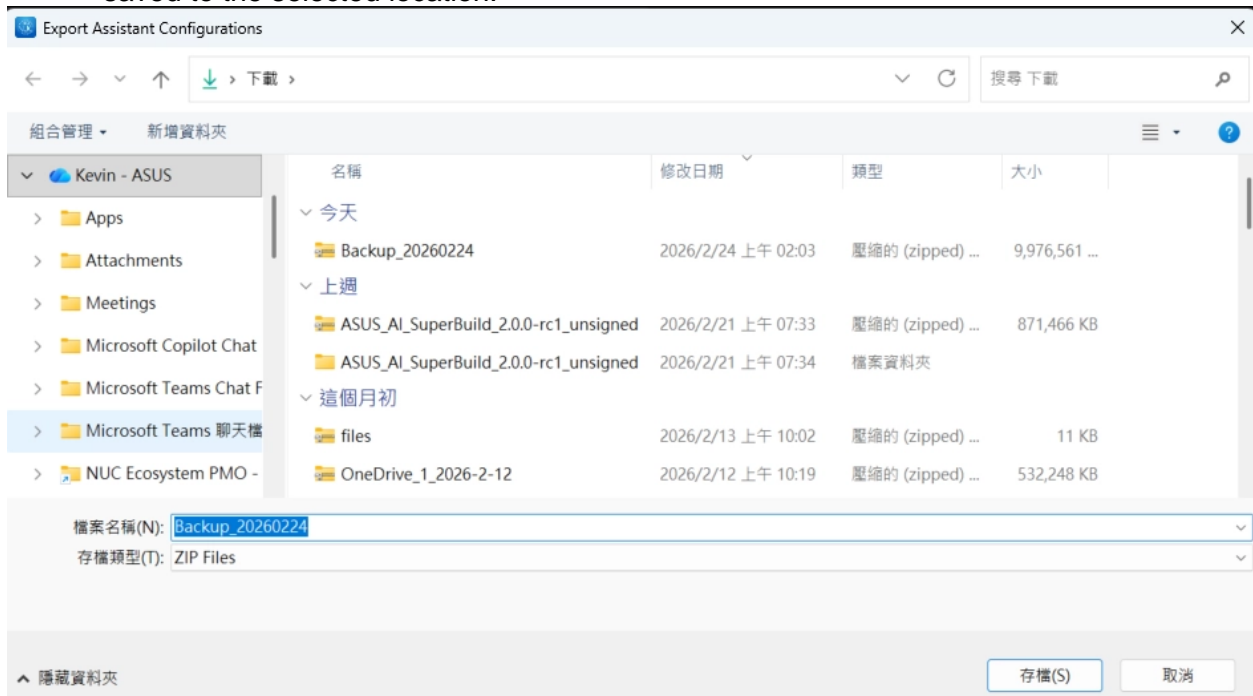


2. Click **“Export Package”** to back up your configuration files. By default, the system will export the following items:
 - **AI Assistant** — all agent configurations
 - **Downloaded Model** — locally downloaded LLM models
 - **Knowledge Base** — RAG knowledge base data



3. Choose a destination folder, then click **“Export”**. The configuration package will be

saved to the selected location.



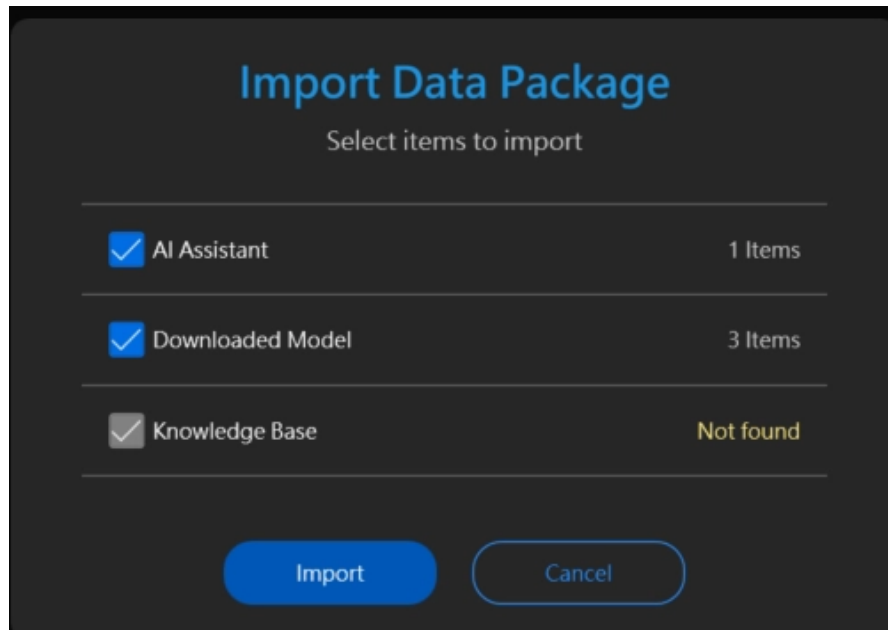
3.7.2 Import Configuration

1. Click “**Import Package**” and select an existing configuration file to restore.

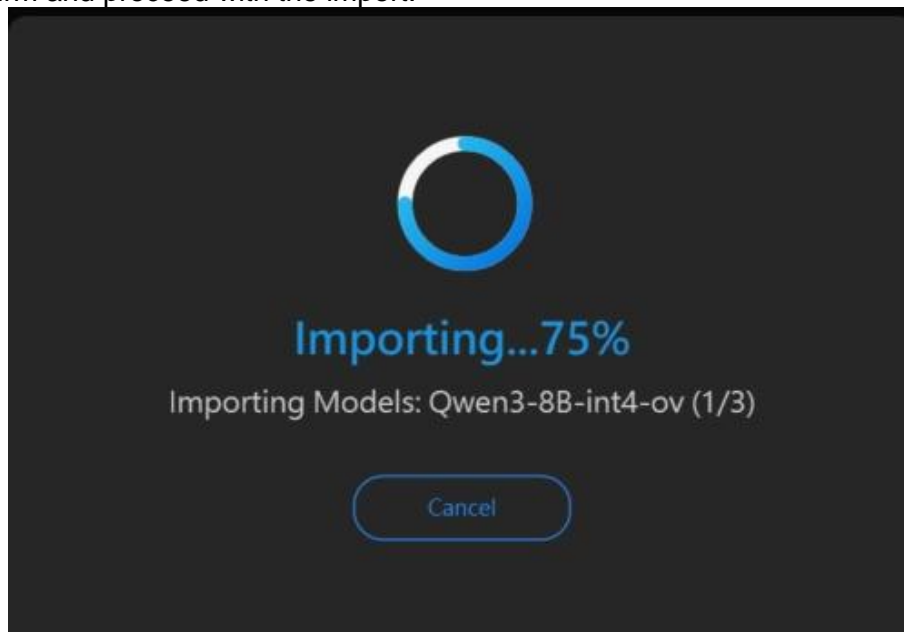
Note: Importing a configuration file will **overwrite** the current data. Please make sure to back up your existing settings before proceeding.



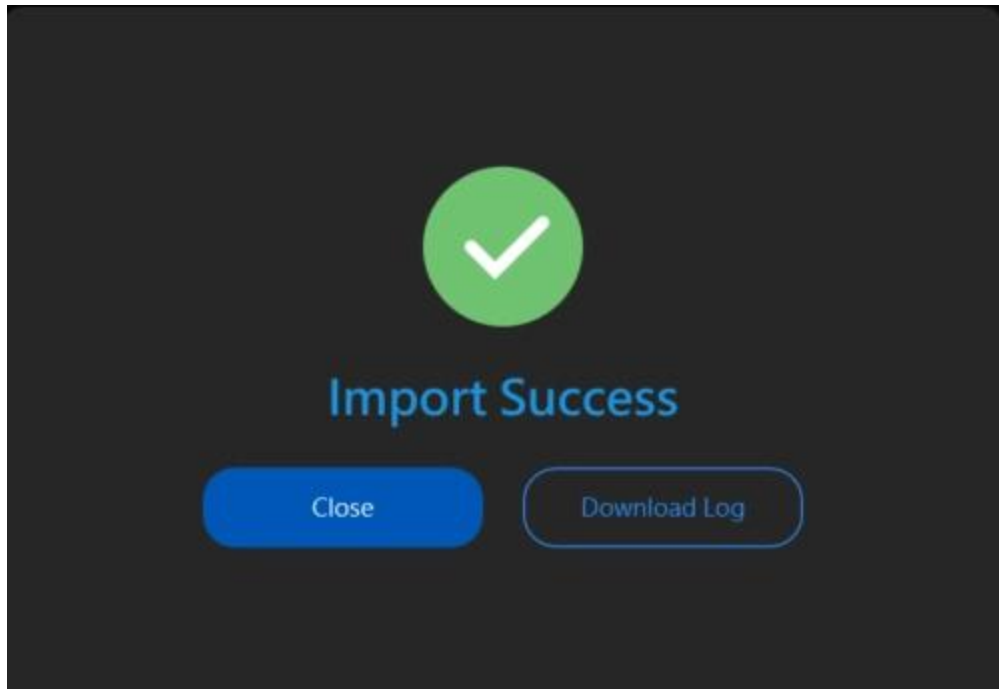
2. After analyzing the file, the system will display a summary of how many configuration items will be overwritten.



3. Confirm and proceed with the import.



4. Once the import is complete, a pop-up window will display the results.

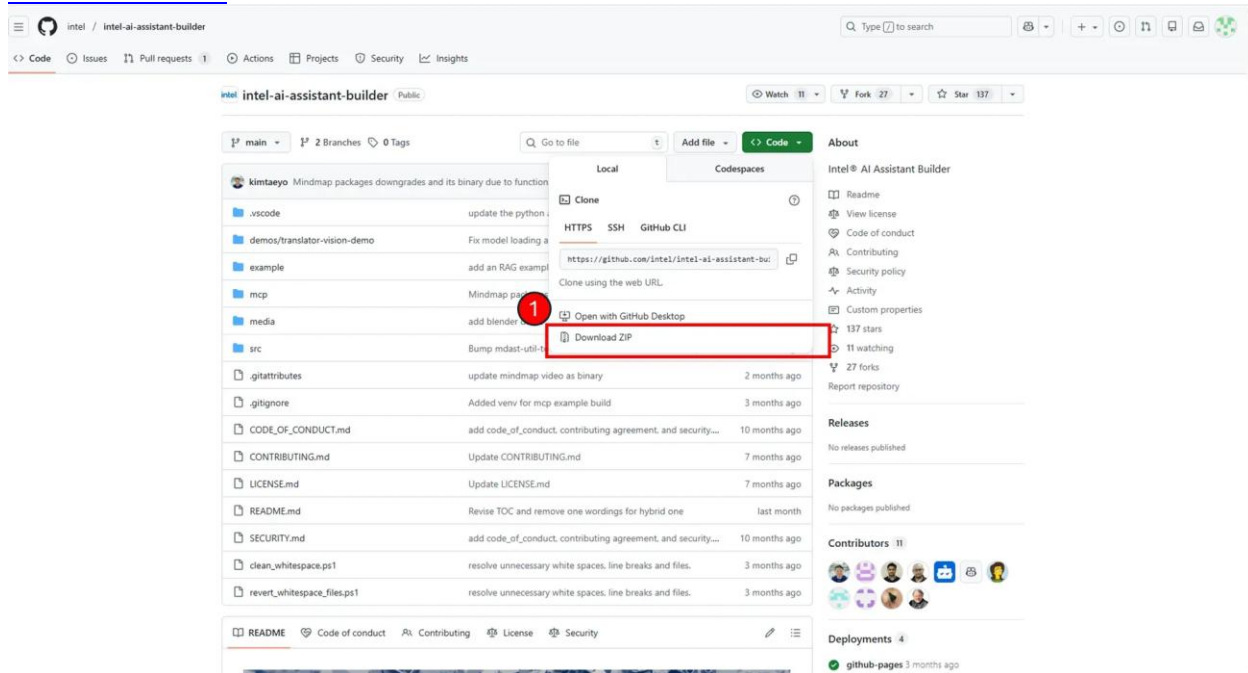


Chapter 4: MCP (Model-Context-Protocol)

You can extend the LLM's capabilities using MCP (Model Context Protocol) by connecting it to your own MCP server to access custom services.

4.1 Example: Build a Hotel Assistant MCP Server

1. Install Python v3.10.x or above.
2. Verify that Python is installed correctly with the command:
python --version
3. Download the example files from GitHub <https://github.com/intel/intel-ai-assistant-builder/tree/main>



4. Open a command line window and navigate to the project folder. (The exact path will be different depending on your environment and where you saved the files.)
cd C:\Users\nuc\Desktop\intel-ai-assistant-builder-main\intel-ai-assistant-builder-main\mcp\mcp_servers\mcp_google_hotel

```
PS C:\Users\nuc> cd C:\Users\nuc\Desktop\intel-ai-assistant-builder-main\intel-ai-assistant-builder-main\mcp\mcp_servers\mcp_google_hotel
```

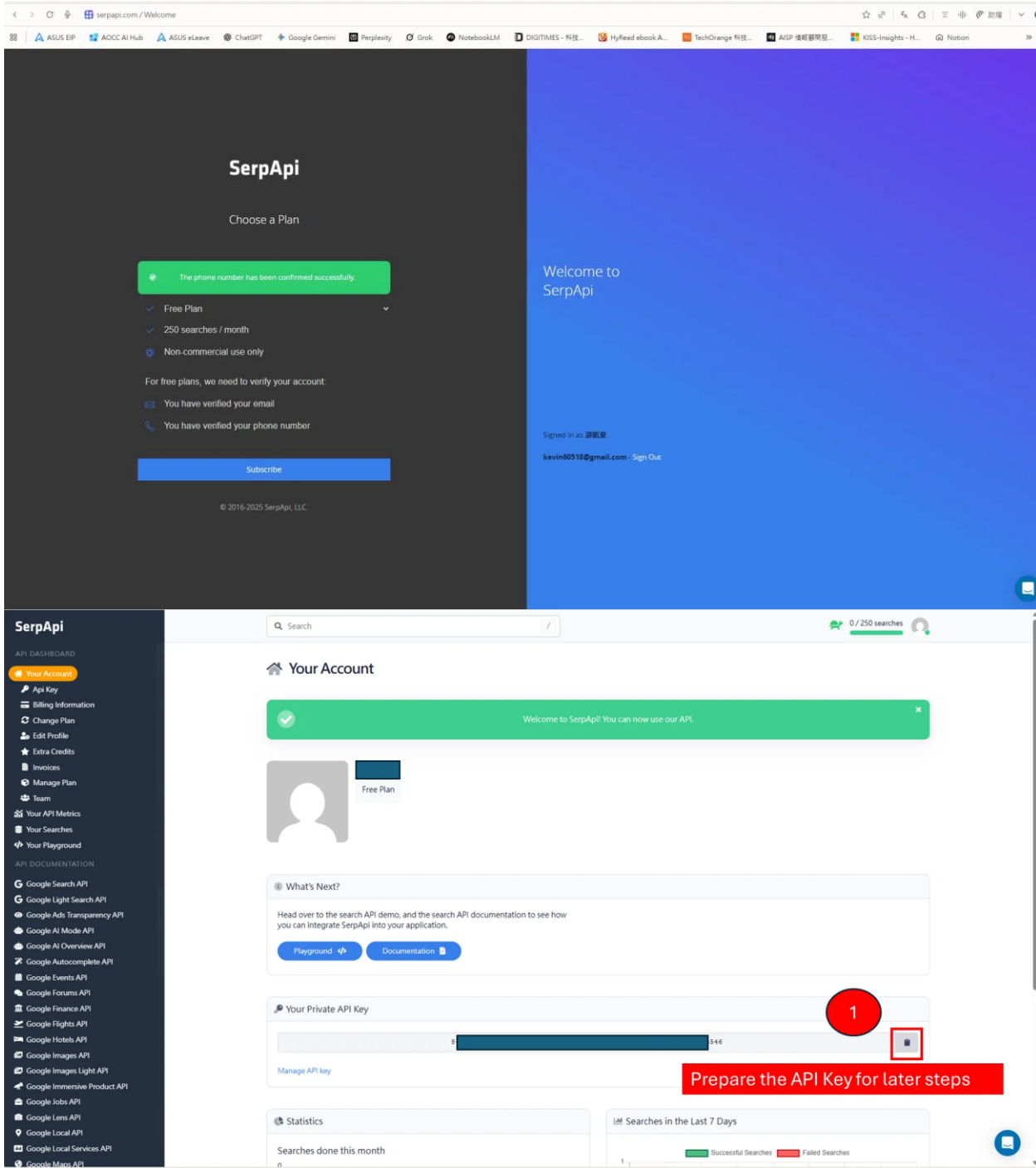
5. Install the dependencies:
python -m pip install -r requirements.txt
6. Build the executable file:
build.bat
7. After the build is complete, you will see the .exe file in the same folder under the dist directory:
C:\Users\nuc\Desktop\intel-ai-assistant-builder-main\intel-ai-assistant-builder-main\mcp\mcp_servers\mcp_google_hotel\dist
8. In **AI SuperBuild**, go to **Assistant Configuration** → **Agent Management** → **MCP**

Server List → Add.

The screenshot shows the ASUS AI SuperBuild interface. On the left sidebar, 'Agent Management' is highlighted with a red box and a red circle containing the number 1. In the main content area, the 'Agent Management' section is active, showing a table with one item: 'Google flight at support flight info'. A red box and a red circle containing the number 2 highlight the 'MCP Server List' tab. Below this, the 'MCP Server List' section is shown, with a red box and a red circle containing the number 3 highlighting the 'Add' button. The table below shows the configuration for the 'Flight MCP' server.

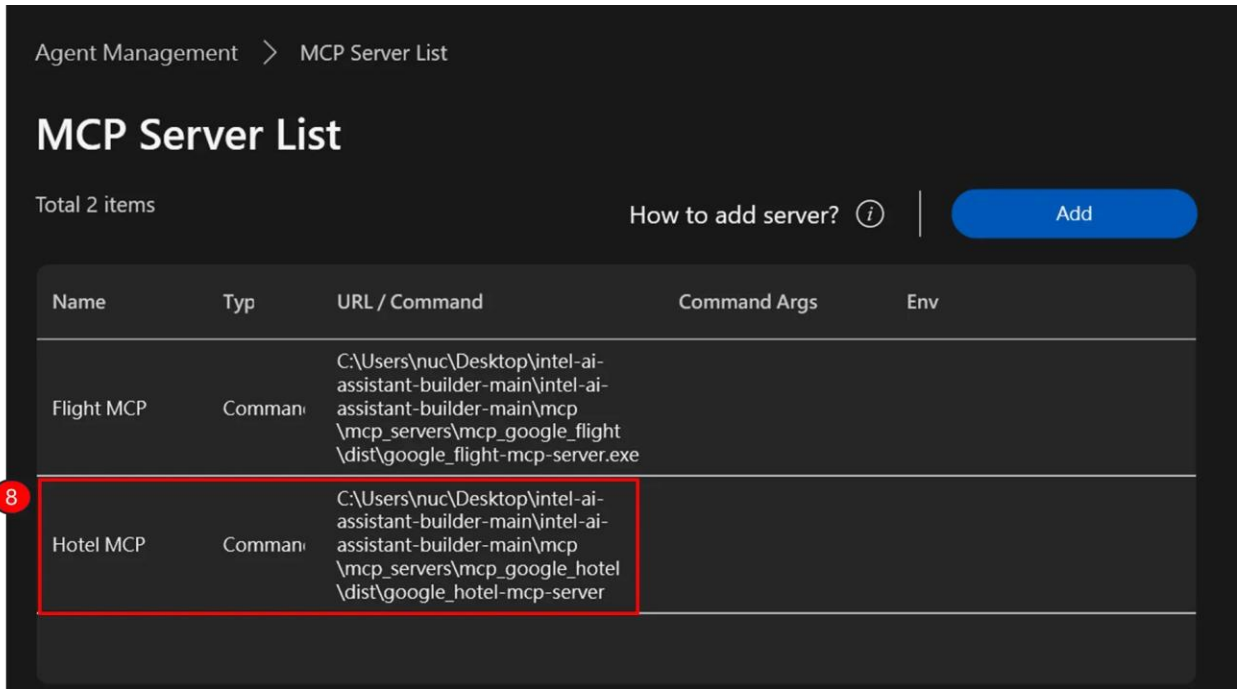
Name	Type	URL / Command	Command Args	Env
Flight MCP	Command	C:\Users\nuc\Desktop\intel-ai-assistant-builder-main\intel-ai-assistant-builder-main\mcp_servers\mcp_google_flight\dist\google_flight-mcp-server.exe		

9. Use the **Command** type to configure the MCP server name and the path to the .exe file. You can get the API Key from the SerpAPI website (https://serpapi.com/users/sign_up).

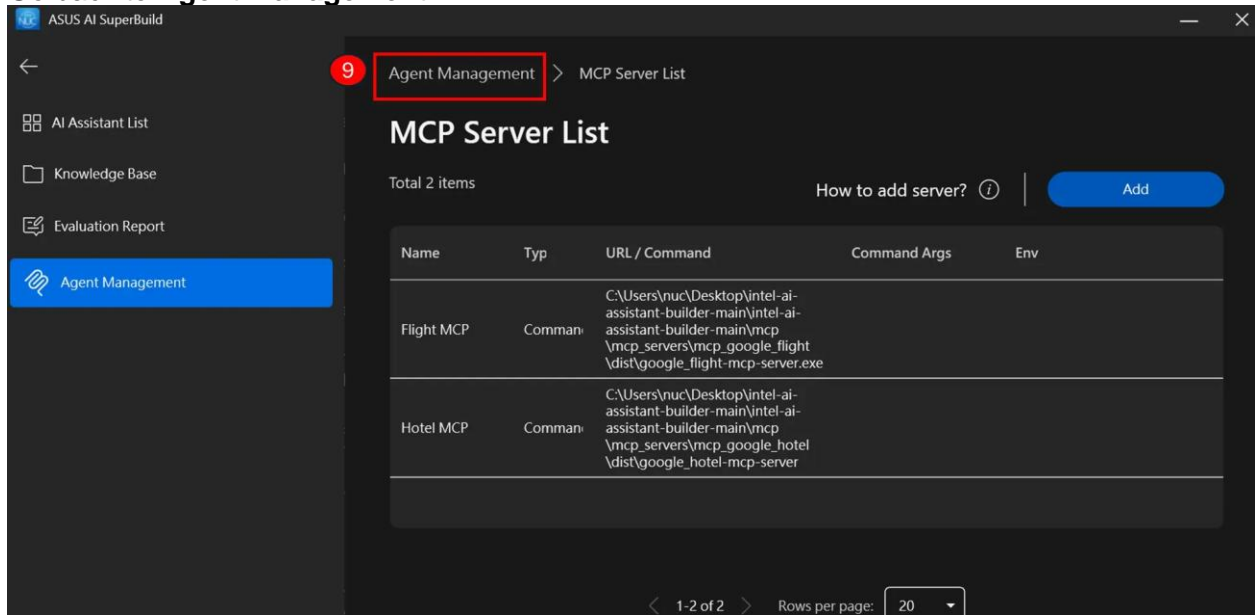


Important: For MCP Server ENV (optional), combine "SERP_API_KEY=" with your API key from SerpAPI. Here's an example:
SERP_API_KEY=8470813f5442b85b22309fc87f65*****546

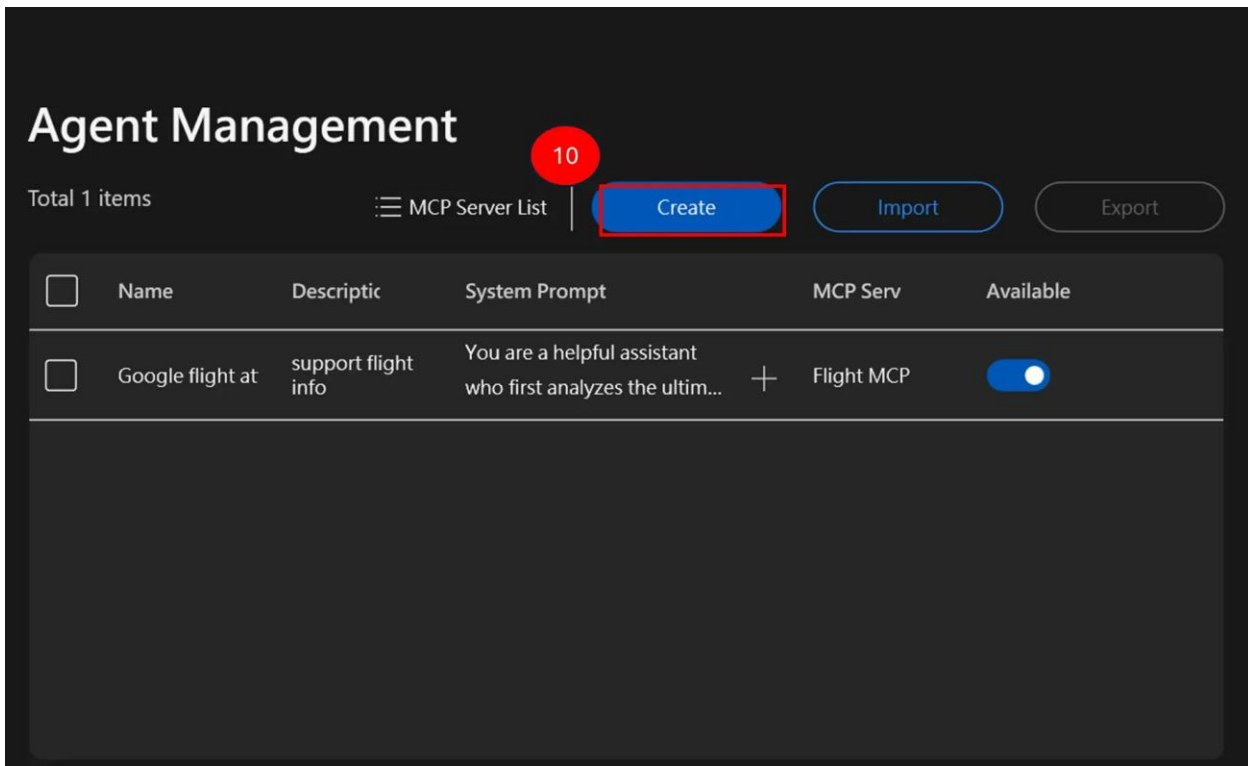
10. The new server will appear in the list.



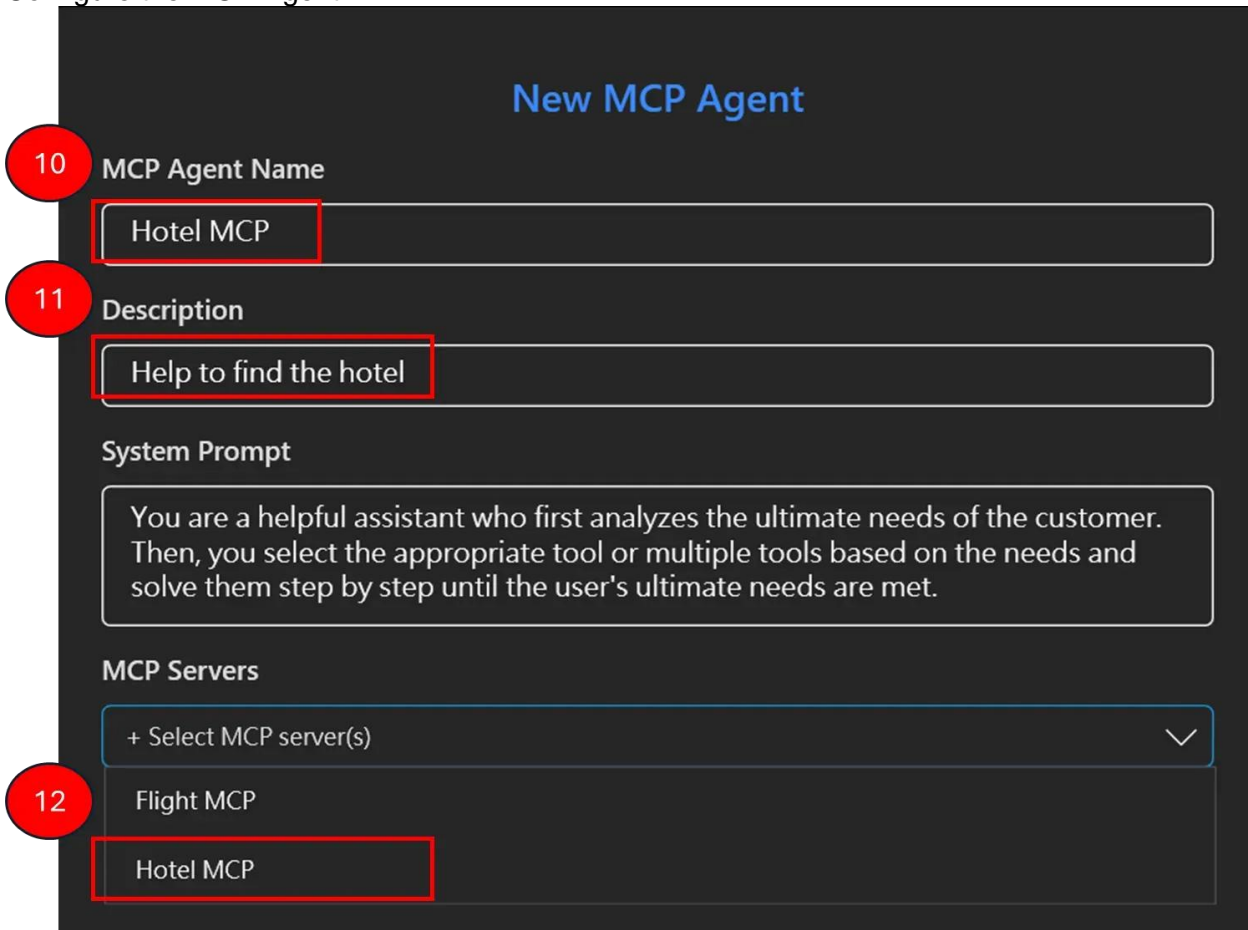
11. Go back to **Agent Management**.



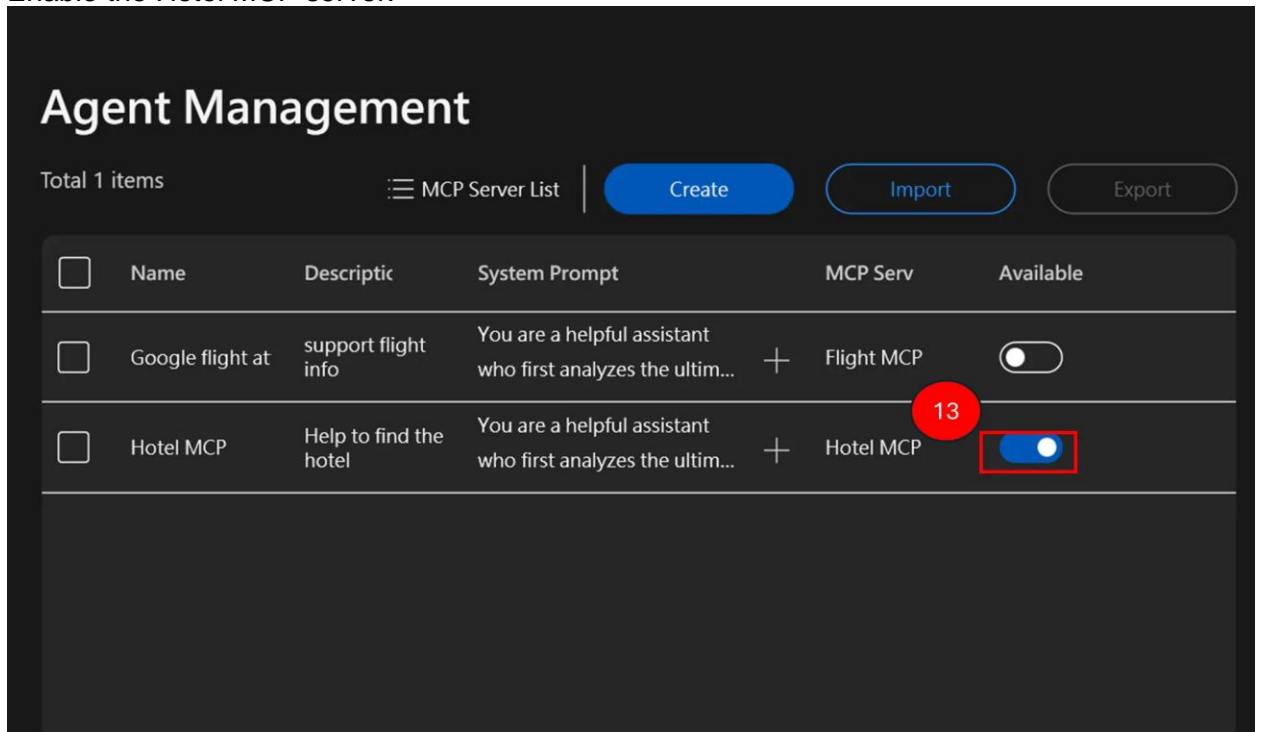
12. Create an MCP Agent task.



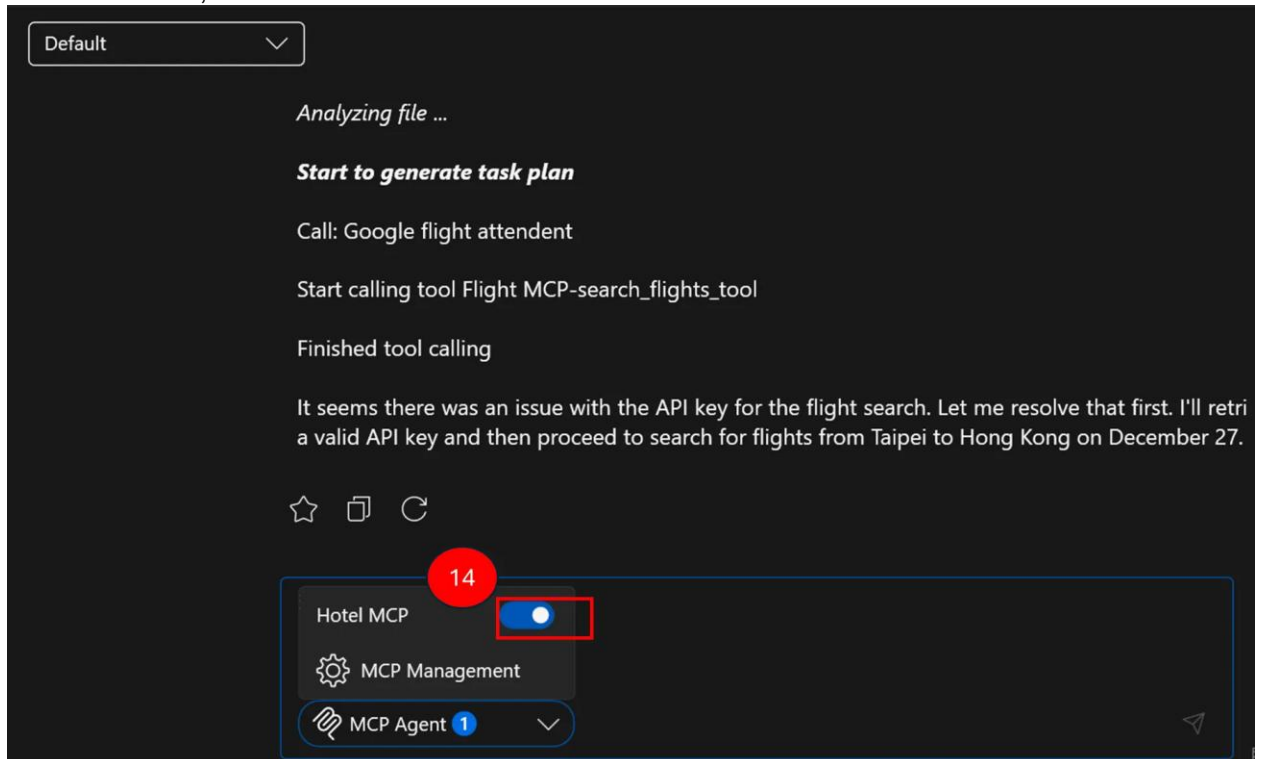
13. Configure the MCP Agent.



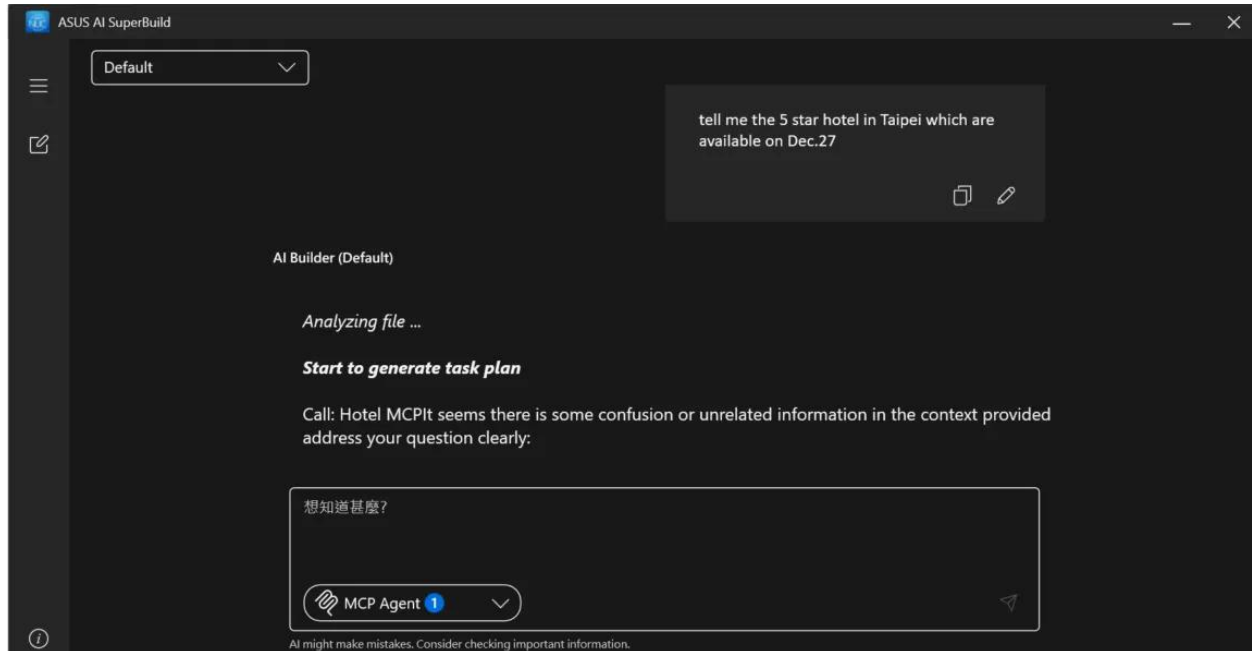
14. Enable the Hotel MCP server.



15. In the chat box, enable the service.



Start chatting and use the MCP tool calling in the conversation.

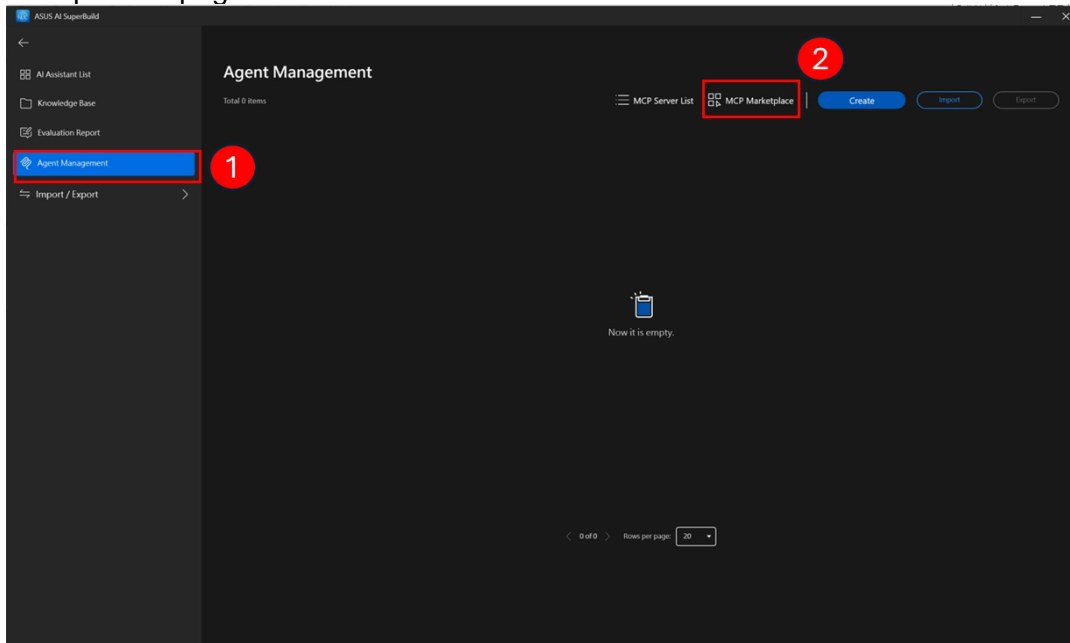


4.2 MCP Marketplace

To help users easily configure MCP servers, AI SuperBuild now supports connecting to MCP with pre-built versions. Navigate to the **MCP Marketplace** to browse and select the MCP tools you need.

1. Navigate to the MCP Marketplace.

Go to **Agent Management** from the left sidebar (1), then click the **MCP Marketplace** tab (2) at the top of the page.



2. Browse and add MCP servers from the Marketplace.

In the MCP Marketplace, you can browse a variety of pre-built MCP servers. Use the

following options to find the right tool:

- **Source from:** Select the source repository — **ModelScope** (default) or **Docker MCP Hub**.
- **Sort by:** Sort the results by popularity (e.g., Viewers High–Low).
- **Search:** Use the search bar to find servers by name, publisher, or keywords.

Each card displays the server name, publisher, description, view count, and category tags. Hover over a card to see more details. Click **"Add to Server List"** to add the desired MCP server to your local server list.

The screenshot displays the MCP Marketplace interface. At the top, there is a breadcrumb trail: "Agent Management > MCP Server List > MCP Marketplace". Below this, the "MCP Marketplace" title is followed by filters: "Source from" set to "ModelScope" and "Sort by" set to "Viewers (High-Low)". A search bar is present with the placeholder text "Search by name, publisher, keywords..." and a "Refresh" button.

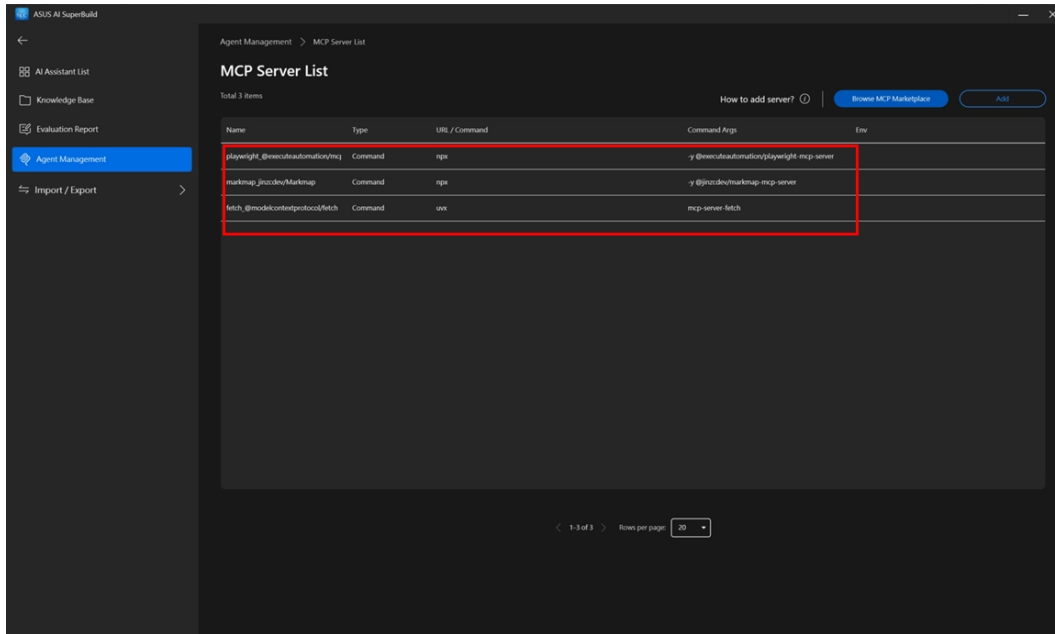
The marketplace contains ten server cards, each with a logo, name, publisher, description, view count, category tags, and an "Add to Server List" button. The cards are:

- Markmap** by jinzcodev/Markmap: # Markmap MCP Server, 965,318 views, tags: knowledge-and-memory, +3.
- leetcode-mcp-server** by @jinzcodev/leetcode-mcp-server: LeetCode MCP server implementation for LeetCode API integration, enabling..., 663,603 views, tags: research-and-data, +4.
- fetch** by @modelcontextprotocol/fetch: This server enables LLMs to retrieve and process content from web pages..., 342,408 views, tag: browser-automation.
- amap-maps** by @amap/amap-maps: Amap Maps is a server that supports any MCP protocol client, allowing..., 222,218 views, tag: location-services.
- 12306-mcp** by @loooooo/12306-mcp: A 12306 ticket search server based on the Model Context Protocol (MCP)..., 117,558 views, tags: travel-and-transportation, search.
- bing-cn-mcp-server** by @yan5236/bing-cn-mcp-server: 让 AI 助手 (如 Claude) 能够使用必应搜索引擎实时获取网络信息的工具。MCP..., 104,978 views, tag: search.
- mcp-server-chart** by @antiviv/mcp-server-chart: # MCP Server Chart ![[0https://badge.mcpx.dev?type=server MCP..., 85,692 views, tags: developer-tools, anti, visualization.
- mcp-server-alipay** by @alipay/mcp-server-alipay: @alipay/mcp-server-alipay 是支付宝开放平台提供的 MCP Server, 让您可以..., 84,695 views, tag: finance.
- sequentialthinking** by @modelcontextprotocol/sequentialthinking: An MCP server implementation that provides a tool for dynamic and..., 51,189 views, tag: research-and-data.
- github** by @modelcontextprotocol/github: MCP Server for the GitHub API, enabling file operations, repository..., 47,649 views, tags: version-control, file-systems, search.

At the bottom of the interface, there is a pagination control showing "1-20 of 100" items and a dropdown menu for "Items per page" set to "20".

3. Verify the added MCP servers.

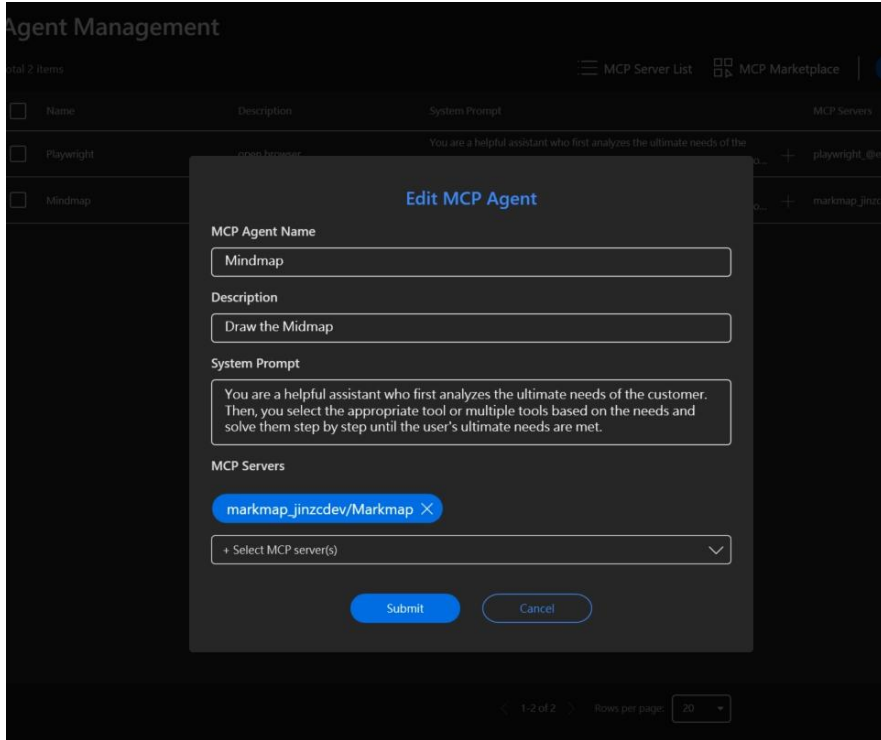
Go back to **Agent Management > MCP Server List** to view the MCP servers you have added. The list shows each server's **Name**, **Type**, **URL / Command**, **Command Args**, and **Env** settings. You can also click **"Browse MCP Marketplace"** to add more servers, or click **"Add"** to manually configure a new server.



4. Create an MCP Agent with the selected servers. Return to the **Agent Management** main page and click **"Create"** to set up a new MCP Agent. In the **Edit MCP Agent** dialog, fill in the following fields:

- **MCP Agent Name:** A descriptive name for the agent (e.g., *Mindmap*).
- **Description:** A brief summary of what this agent does (e.g., *Draw the Mindmap*).
- **System Prompt:** Define the agent's behavior and instructions.
- **MCP Servers:** Select one or more MCP servers from the dropdown to assign to this agent.

Click **"Submit"** to save. The agent will now be able to leverage the selected MCP tools to accomplish complex, multi-step tasks.

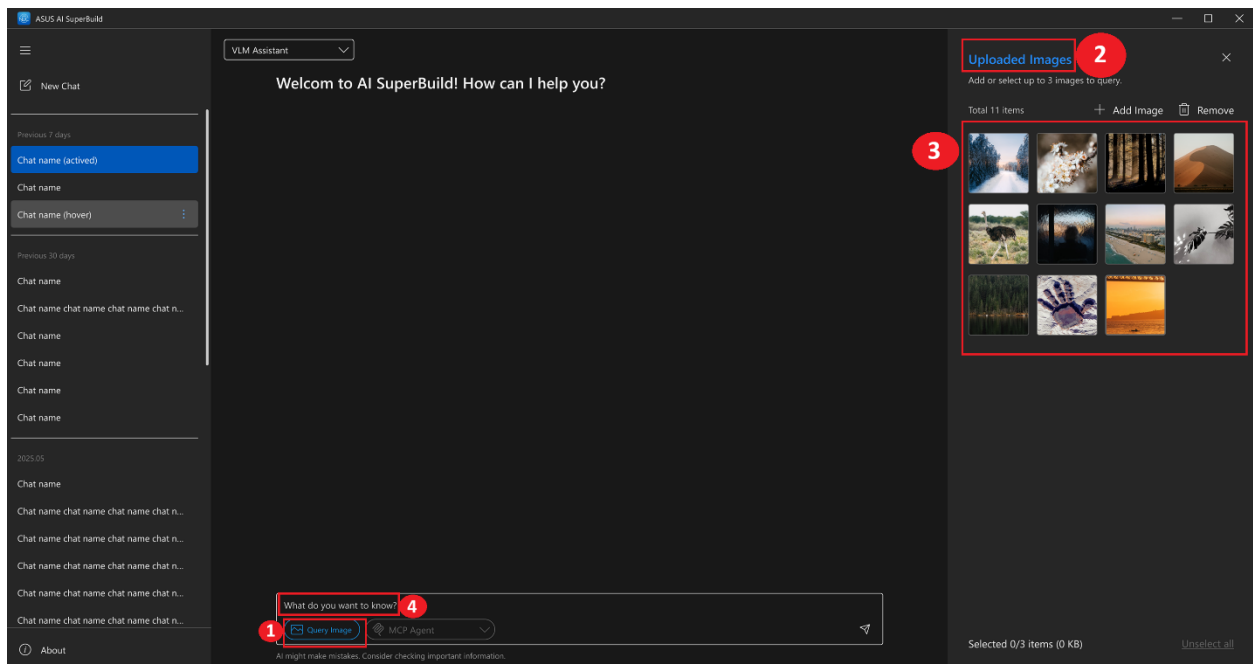


Chapter 5: Vision Language Model(Image Analysis)

You can easily utilize Vision Language Models (VLM) directly within the main chat interface. Simply upload any image or paste a screenshot, and will automatically analyze its contents to assist with your LLM queries.

5.1 Executing VLM in the Main Chat

1. **Initiate Image Query:** In the Main Chat interface, select the Query Image option.
2. **Upload Your File:** Click on Upload Image to browse your device and upload the picture or screenshot you wish to analyze.
3. **Select the Image:** Once uploaded, click on the specific image you want to use from the image list.
4. **Ask Your Question:** Type any question regarding the selected image directly into the Main Chat input box and press Enter to receive the AI's analysis.



Appendix

1. MCP Example Setting SOP

1.1 MCP Environment Setup

This guide provides step-by-step instructions to prepare your Windows system. Completing these prerequisite steps ensures that the ASUS NUC AI Super Build MCP scenarios will run smoothly during your demonstrations.

⚠ Important Notice Before Starting:

Please save all your ongoing work and close other applications. Your computer will automatically restart twice during this installation process to properly configure the necessary system environment variables in the background.

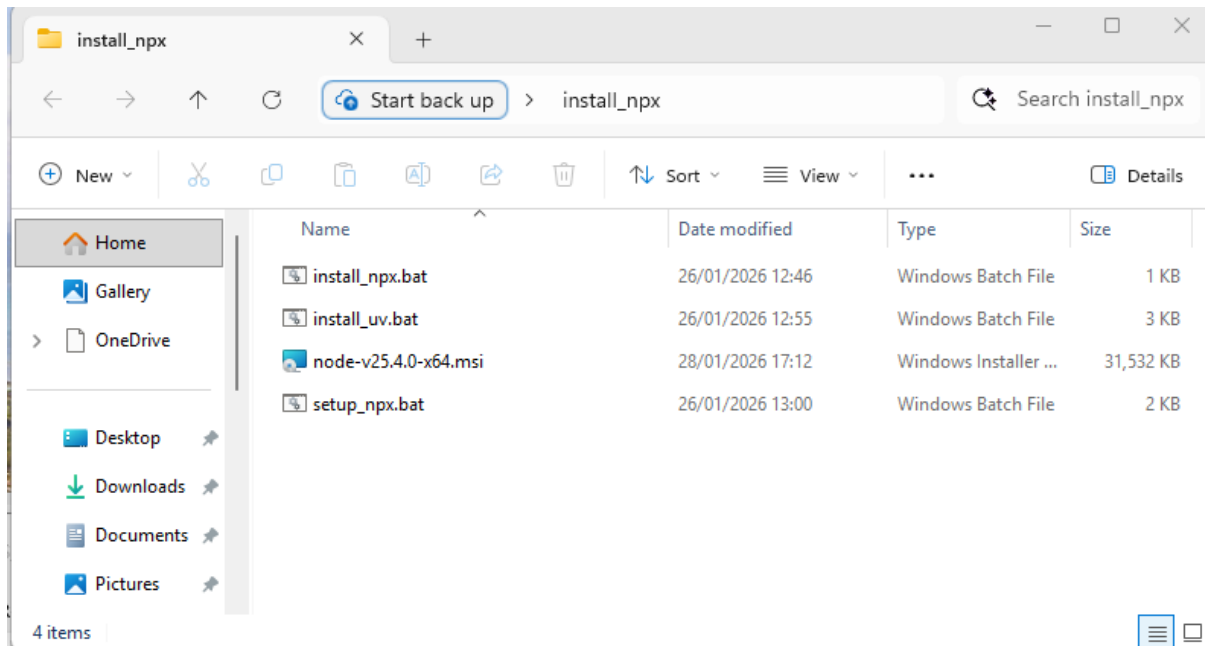
A. Phase 1: Preparation

- a. Download the file: [install_npx.zip](#)
- b. Locate the File: Find the compressed installation file named install_npx.zip provided by the team.



- c. Extract the Files: Right-click the compressed file and select Extract All... to unzip its contents into a standard folder.

- d. Open the Folder: Open the newly extracted folder. You will find three key executable files inside: `install_npx`, `setup_npx`, and `install_uv`.

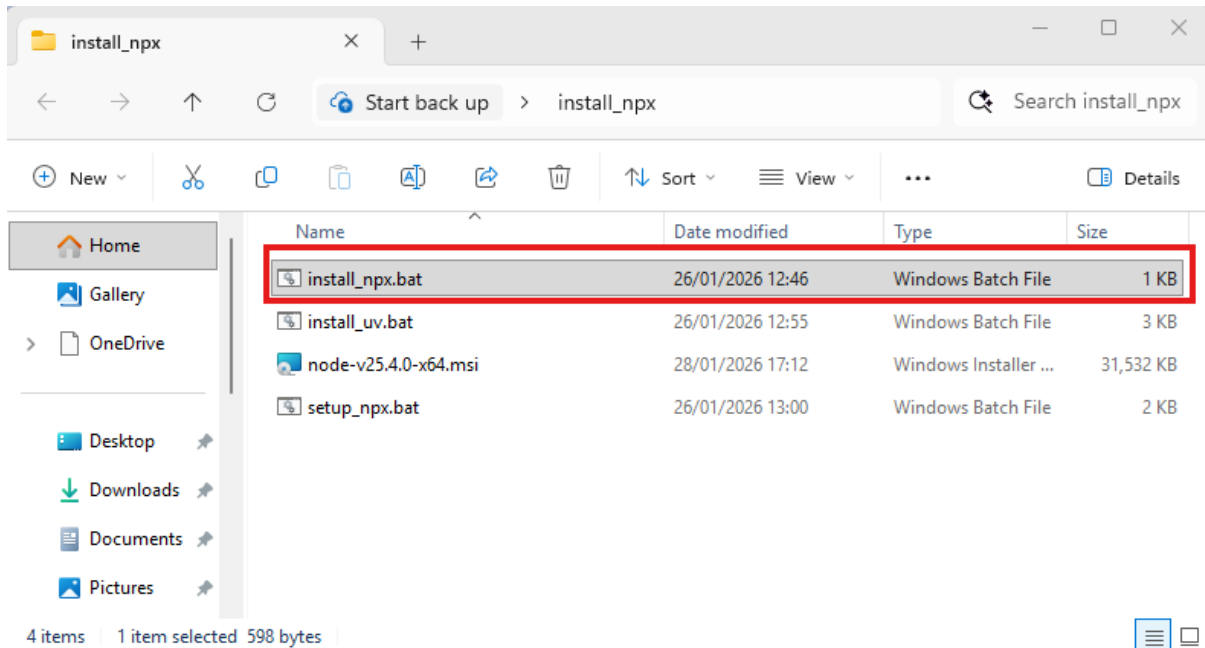


B. Phase 2: Automated Installation & Configuration

The team has automated the complex environment variable settings. Please simply run the three files in the exact order below.

a. Step 1: Run install_npx.bat

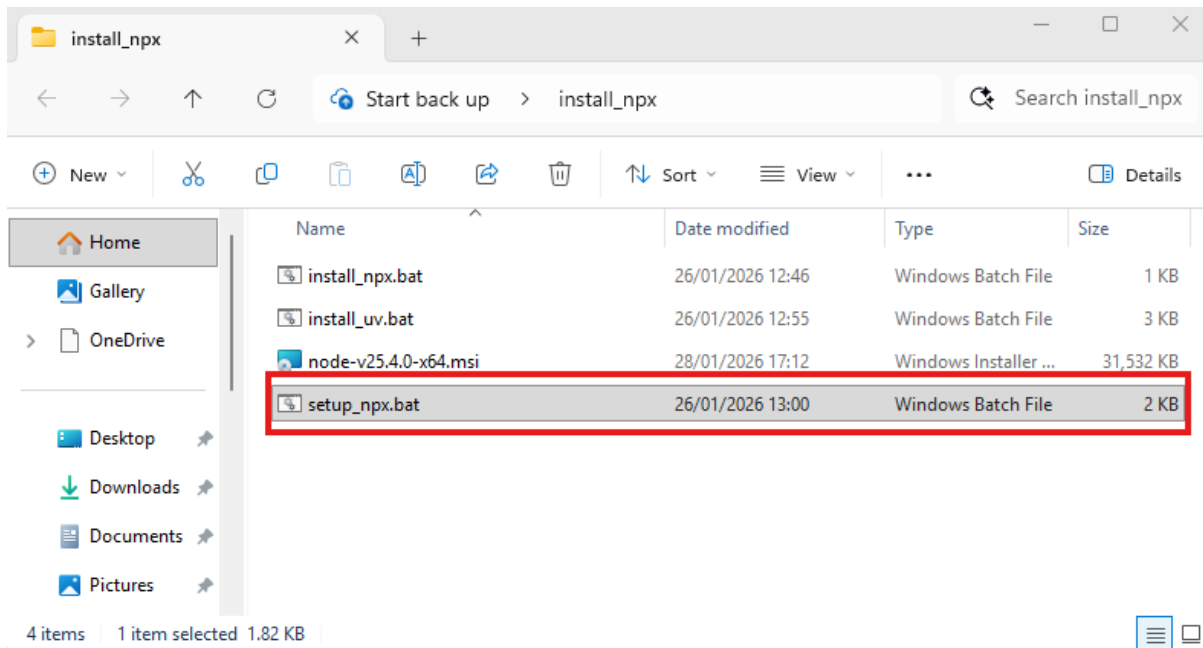
- Double-click the install_npx.bat file.



- A command window will appear while it automatically sets up the initial environment variables. Wait for the process to finish and the window to close.

b. Step 2: Run setup_npx .bat (First Reboot)

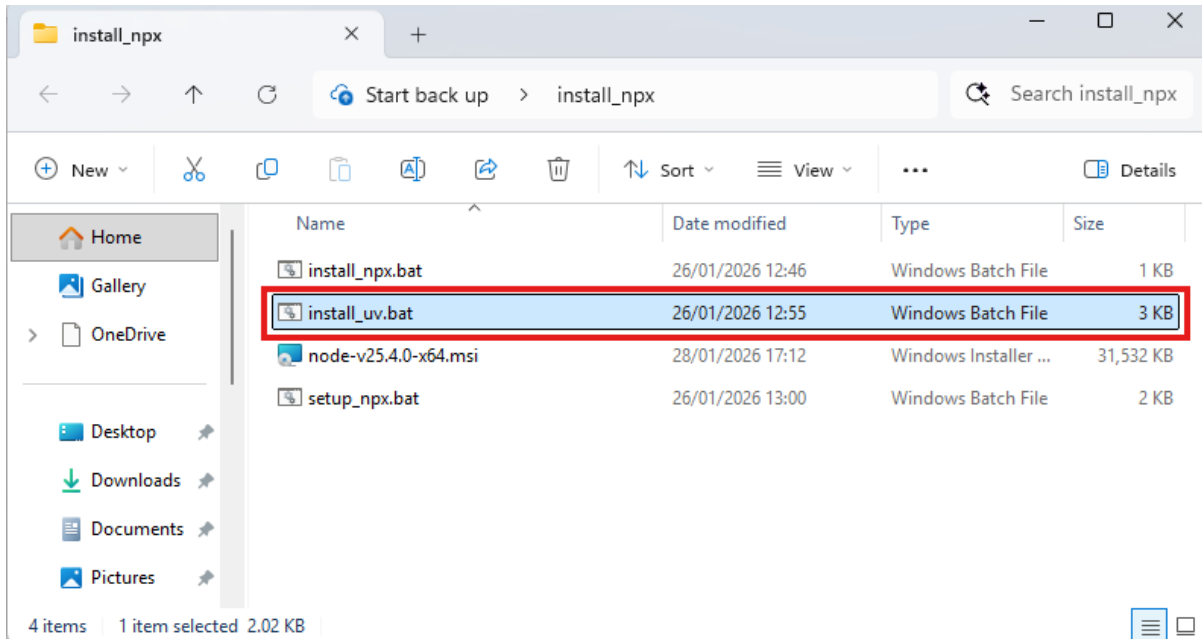
- Double-click the setup_npx.bat file.



- Allow the script to run its course.
- Note: Once this process completes, your NUC will automatically restart to apply the first set of environment variables.
- After your PC reboots, log back into Windows and re-open the extracted folder to continue to the final step.

c. Step 3: Run install_uv.bat (Second Reboot)

- Double-click the install_uv.bat file.

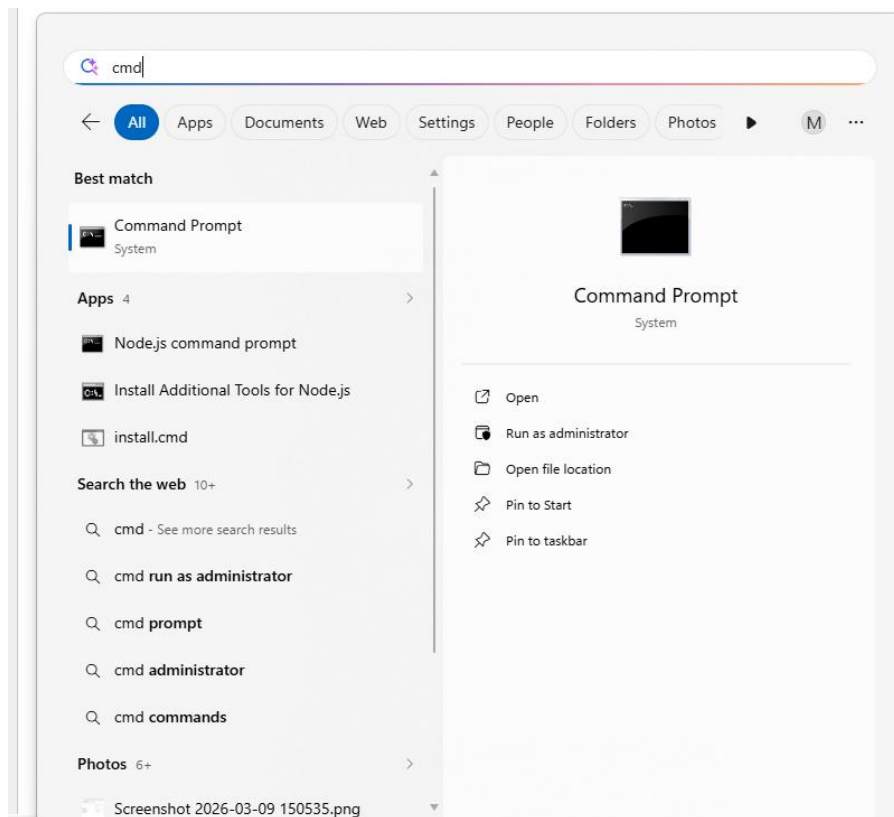


- Wait for the installation and environment variable setup to finish.
- Note: Your computer will automatically restart one last time to finalize all system configurations.
- Log back into Windows. The automated setup is now complete!

C. Phase 3: Verifying the Setup

Let's make sure everything is installed correctly and communicating with the system before you start your demo.

- a. Open Command Prompt: Press the Windows Key on your keyboard, type cmd, and press Enter. A black text window will open.

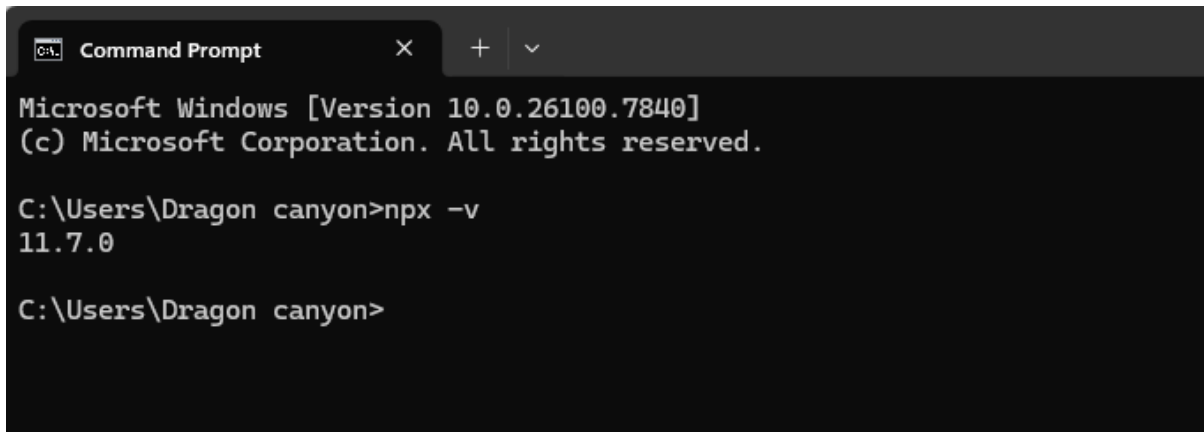


- b. Check npx: Type npx -v and press Enter.

```
Command Prompt
Microsoft Windows [Version 10.0.26100.7840]
(c) Microsoft Corporation. All rights reserved.

C:\Users\Dragon canyon>npx -v
```

- c. Success: You will see a version number appear on the next line.

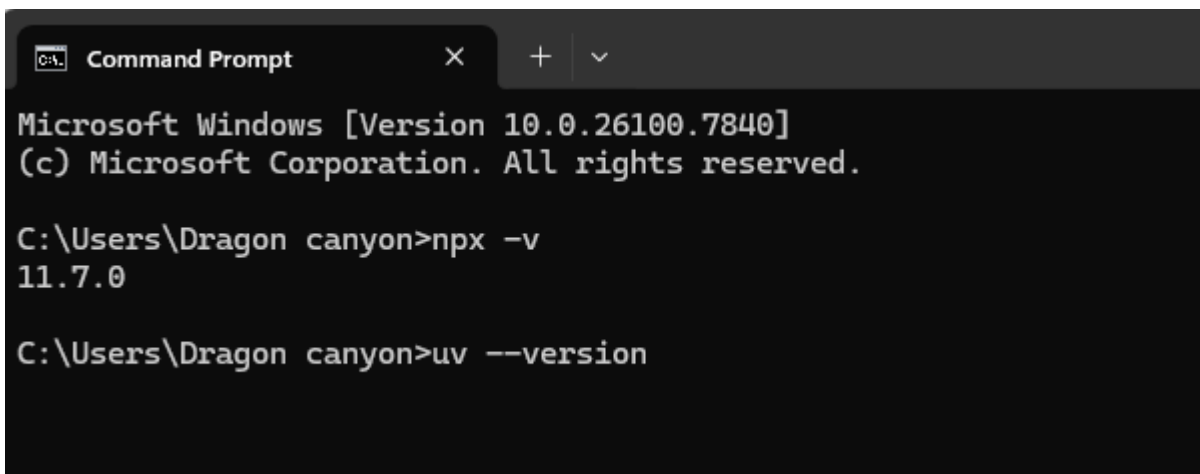


```
Command Prompt
Microsoft Windows [Version 10.0.26100.7840]
(c) Microsoft Corporation. All rights reserved.

C:\Users\Dragon canyon>npx -v
11.7.0

C:\Users\Dragon canyon>
```

- d. Check uv: Type `uv --version` and press Enter.

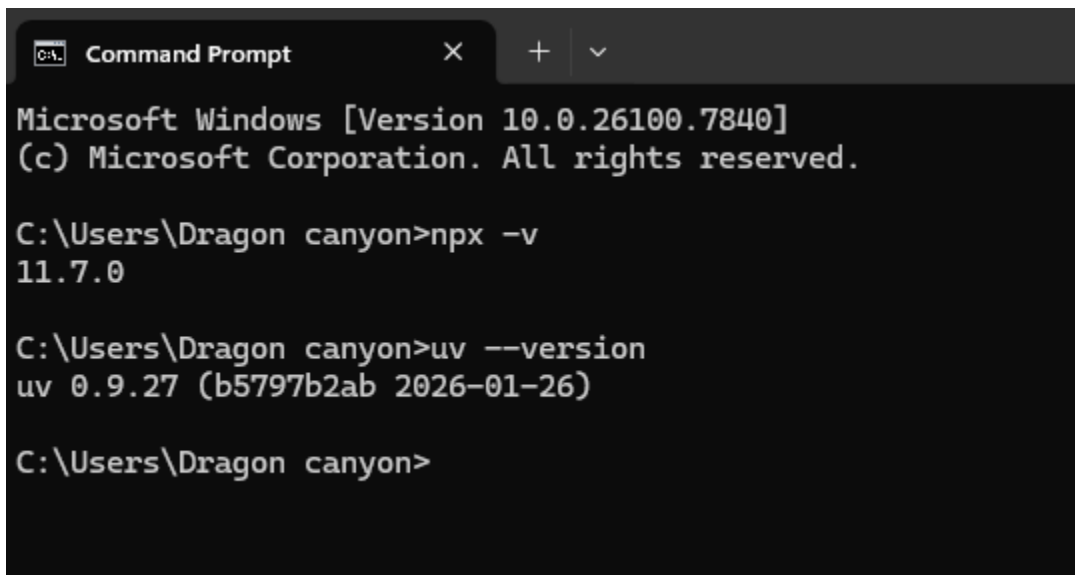


```
Command Prompt
Microsoft Windows [Version 10.0.26100.7840]
(c) Microsoft Corporation. All rights reserved.

C:\Users\Dragon canyon>npx -v
11.7.0

C:\Users\Dragon canyon>uv --version
```

- e. Success: You will see a version number appear.



```
Command Prompt
Microsoft Windows [Version 10.0.26100.7840]
(c) Microsoft Corporation. All rights reserved.

C:\Users\Dragon canyon>npx -v
11.7.0

C:\Users\Dragon canyon>uv --version
uv 0.9.27 (b5797b2ab 2026-01-26)

C:\Users\Dragon canyon>
```

- f. Close the Window: If both commands successfully returned version numbers, your environment setup is 100% ready. You can close the command prompt and proceed with the MCP demo scenarios.

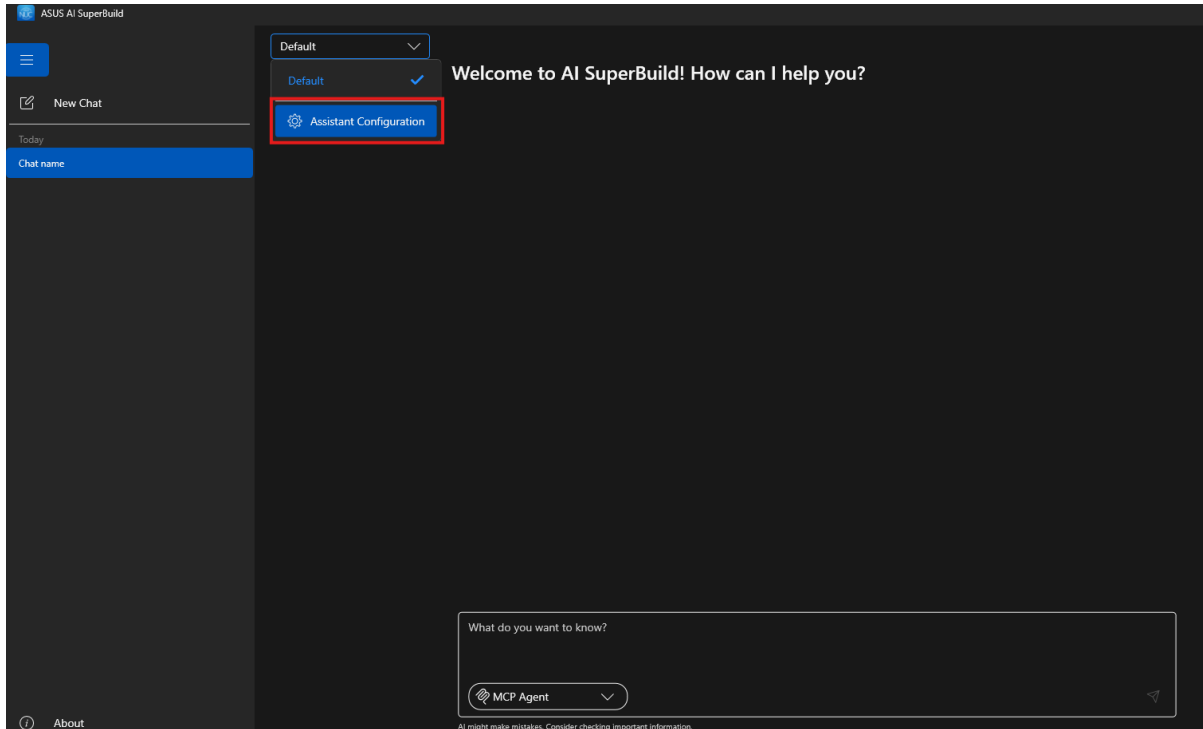
2.1 MCP Example Scenarios Setup SOP

A. Scenario 1: Product Information Retrieval (Knowledge Base Setup)

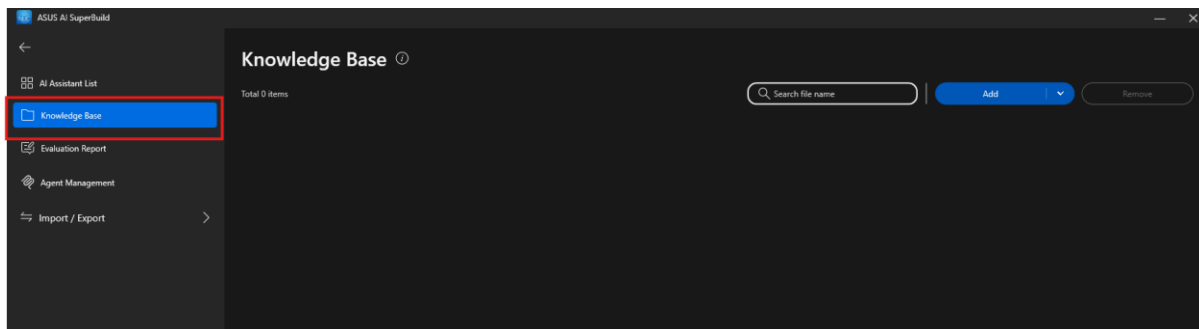
This scenario demonstrates the AI's ability to learn from uploaded documents and accurately answer specific product queries. Please note that this initial scenario uses the built-in Knowledge Base feature rather than the MCP tools.

Step-by-Step Instructions:

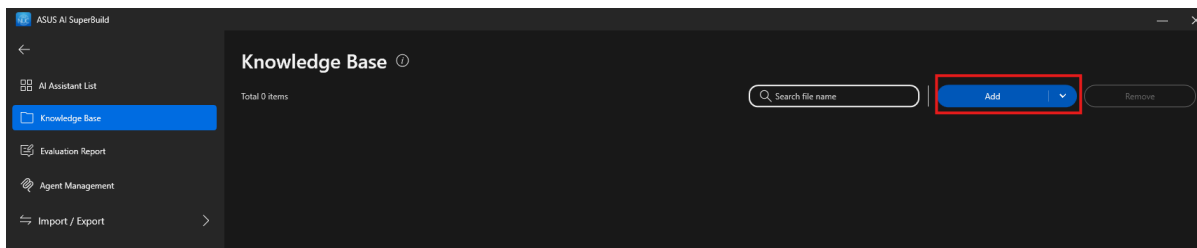
A. Access Settings: On the main application interface, click on Assistant Configuration.



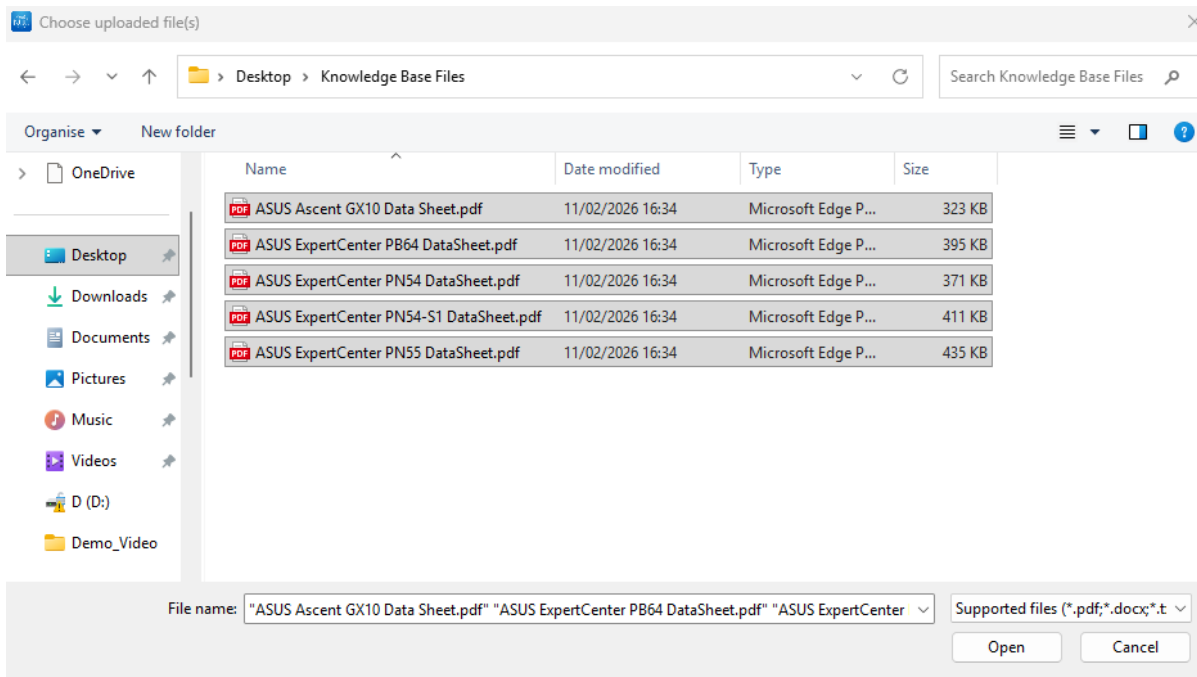
B. Open Knowledge Base: From the configuration menu, select Knowledge Base.



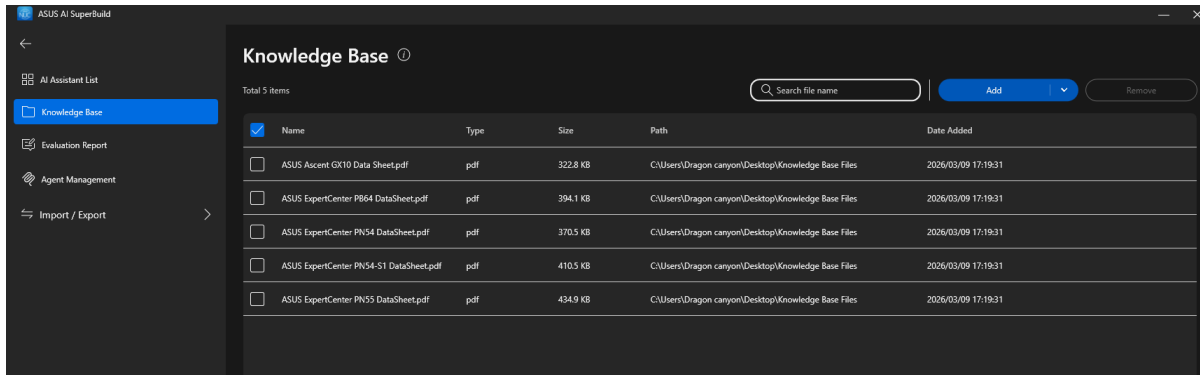
C. Add Document: Click the Add button to upload new reference materials.



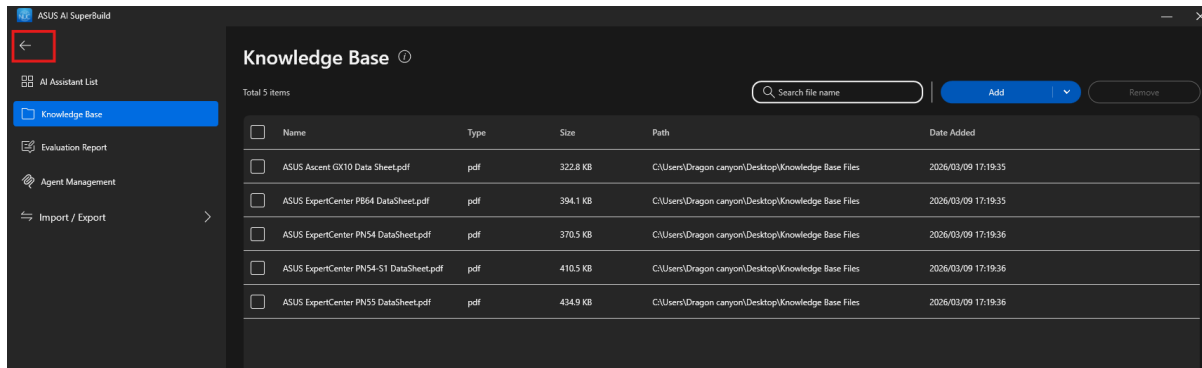
D. Upload Files: Select the provided ASUS product information documents from your computer and upload them.



E. Wait for Processing: Please wait a few moments until the document upload is fully completed and processed by the system.

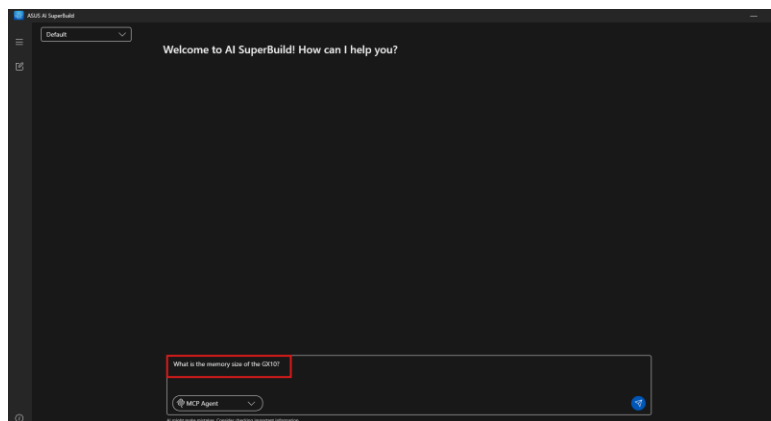


F. Return to Home: Click the Back arrow to return to the main chat interface.



G. Start the Demo: You can now demonstrate the AI's capability by asking specific questions about the uploaded product.

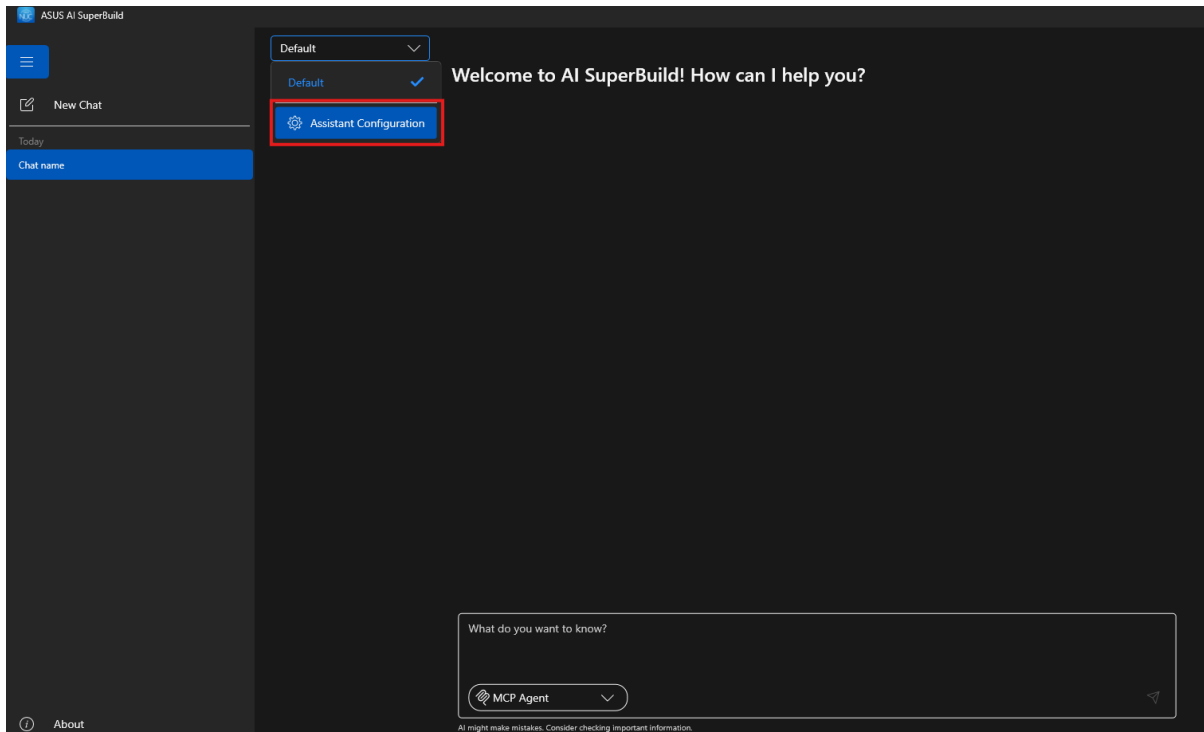
H. Demo Question Example: Type "What is the memory size of the GX10?" and press Enter to see the AI retrieve the exact specification directly from your uploaded documents.



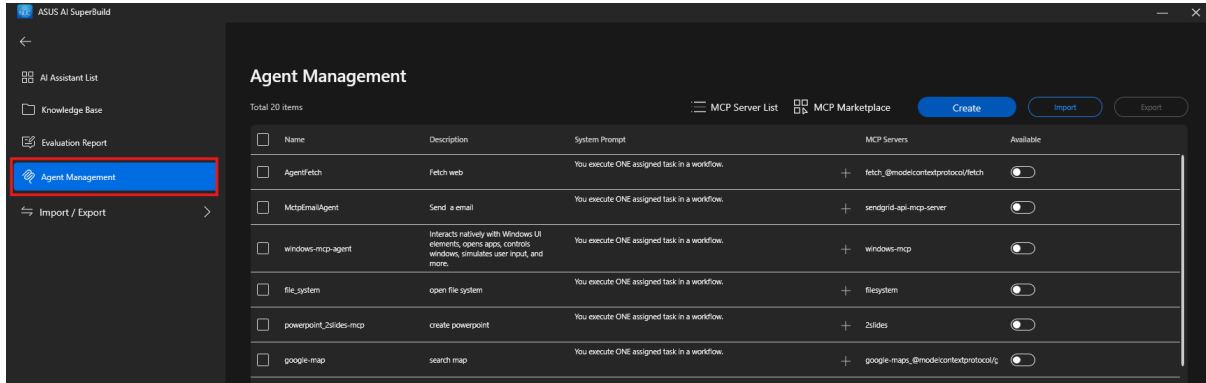
B. Scenario 2: Real-Time Competitive Analysis (Brave Search MCP Setup)

This scenario demonstrates how to configure AI Super Build to browse the live internet using the Brave Search MCP. This allows the sales team to perform real-time competitive analysis during a customer demonstration.

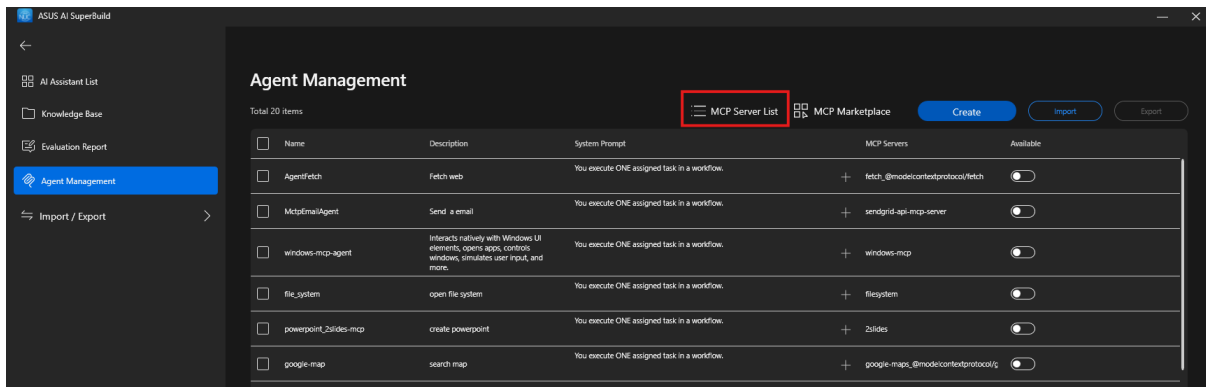
- A. Step 1: Obtain a Brave Search API Key (Prerequisite)
 - a. Before configuring the system, you must have an active API key from Brave.
 - b. Go to the official [Brave Search API website] and register for an account.
 - c. Generate and copy your unique API Key. Keep this handy for Step 3.
- B. Step 2: Add the MCP Server
 - a. On the main application interface, click on Assistant Configuration.



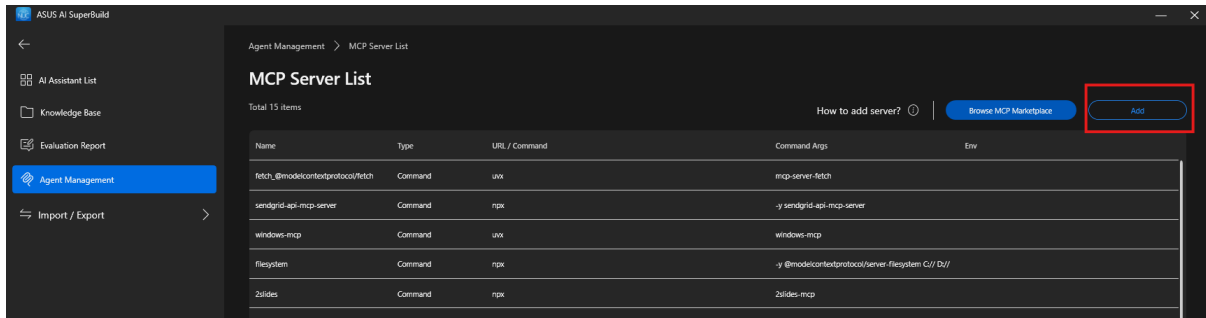
b. Select Agent Management.



c. Click on the MCP Server List tab.



d. Click the Add button to configure a new server.



- e. In the configuration window, set the Type dropdown menu to Command.

New MCP Server

Type

URL **Command**

MCP Server Name

MCP Server Command

MCP Server Command Args (optional)

MCP Server Env (optional)

e.g. API_KEY

Add Cancel

C. Step 3: Configure the Server Details

Fill out the form using the exact text provided below.

- MCP Server Name: brave-search
- MCP Server Command: npx
- MCP Server Command Args (optional): -y @brave-search-mcp-server --transport stdio
- MCP Server Env (optional): BRAVE_API_KEY: [Paste Your API Key Here] (Make sure there are no extra spaces before or after the key)
- Click Submit to save the server configuration.

Edit MCP Server

Type

URL Command

MCP Server Name

brave-search

MCP Server Command

npx

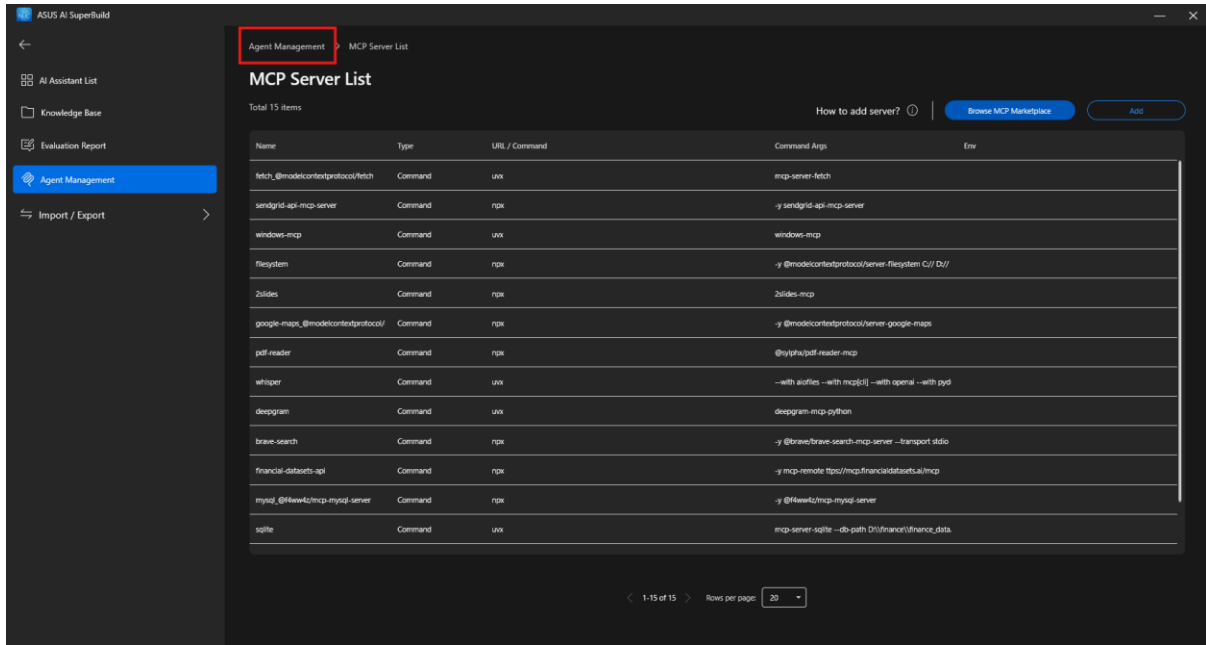
MCP Server Command Args (optional)

-y @brave/brave-search-mcp-server --transport stdio

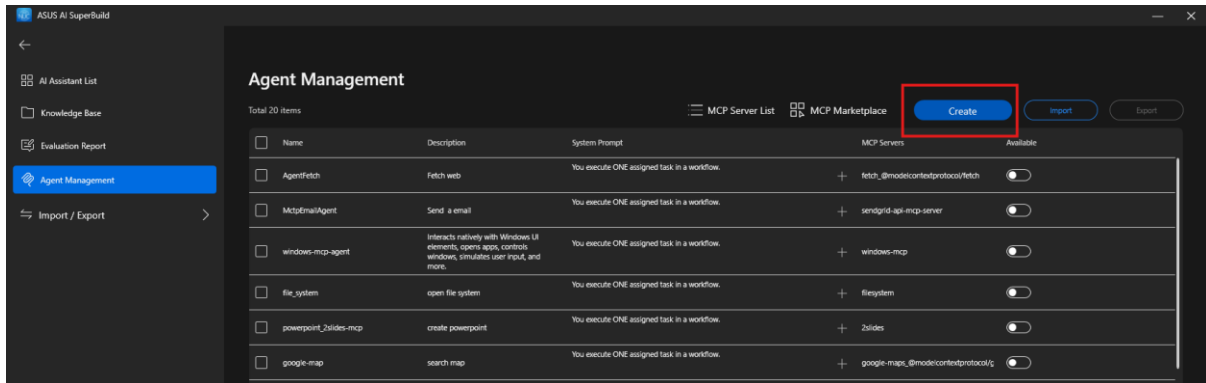
MCP Server Env (optional)

Submit Cancel

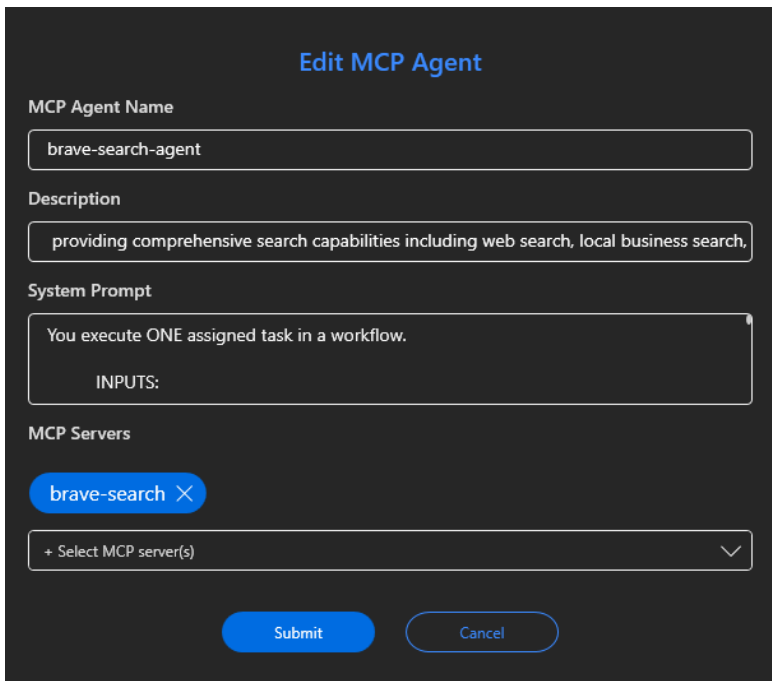
- D. Step 4: Create the MCP Agent
 - a. Return to the main Agent Management page.



- b. Click the Create button.

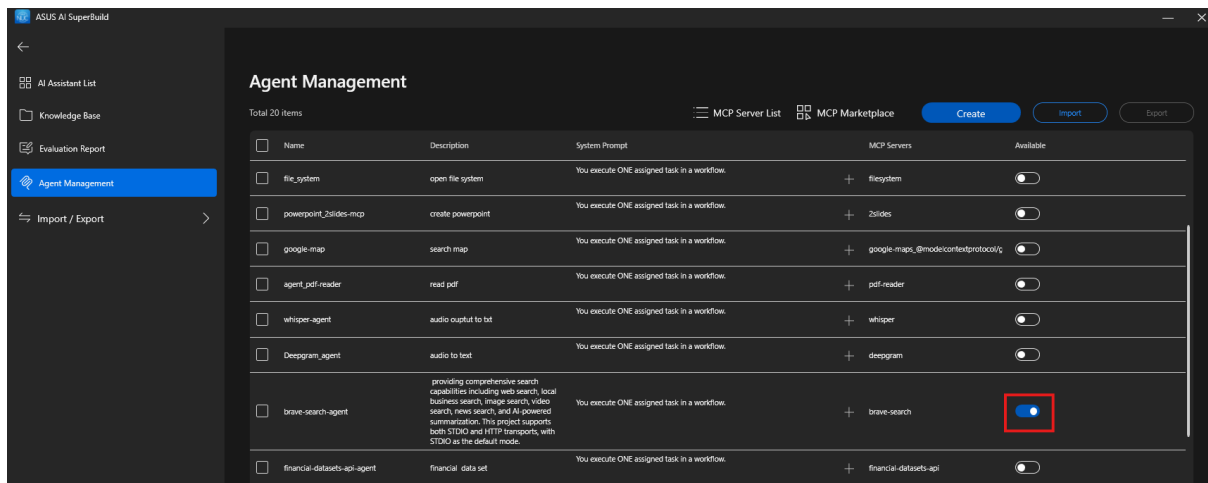


- c. Fill out the Agent details as follows:
 - a. MCP Agent Name: brave-search-agent
 - b. Description: search
 - c. System Prompt: (no modification is needed)
 - d. MCP Servers: Check the box for brave-search (the server you just created in Step 3).
- d. Click Submit.

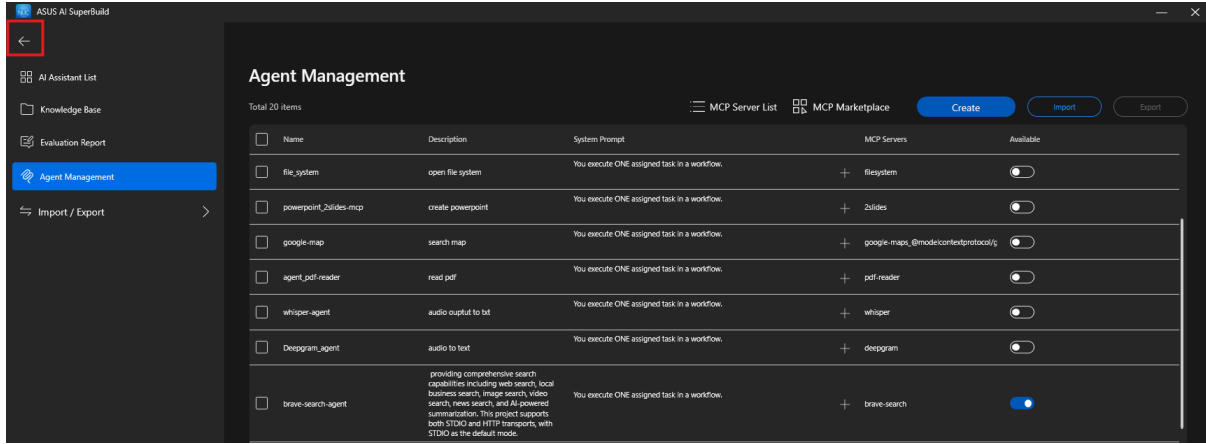


E. Step 5: Enable and Execute the Demo

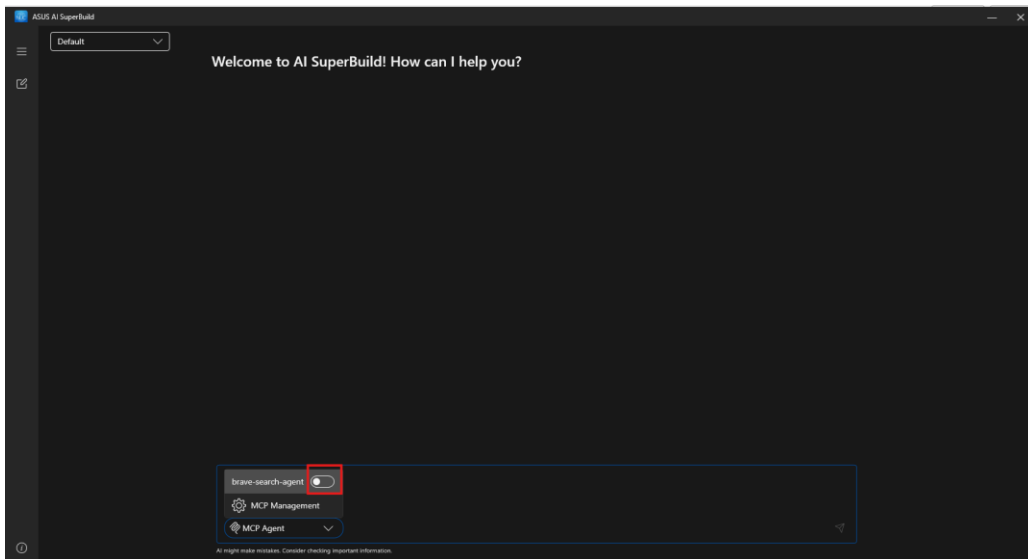
- a. Locate your newly created brave-search-agent in the list and toggle the switch to Enable Brave Search.



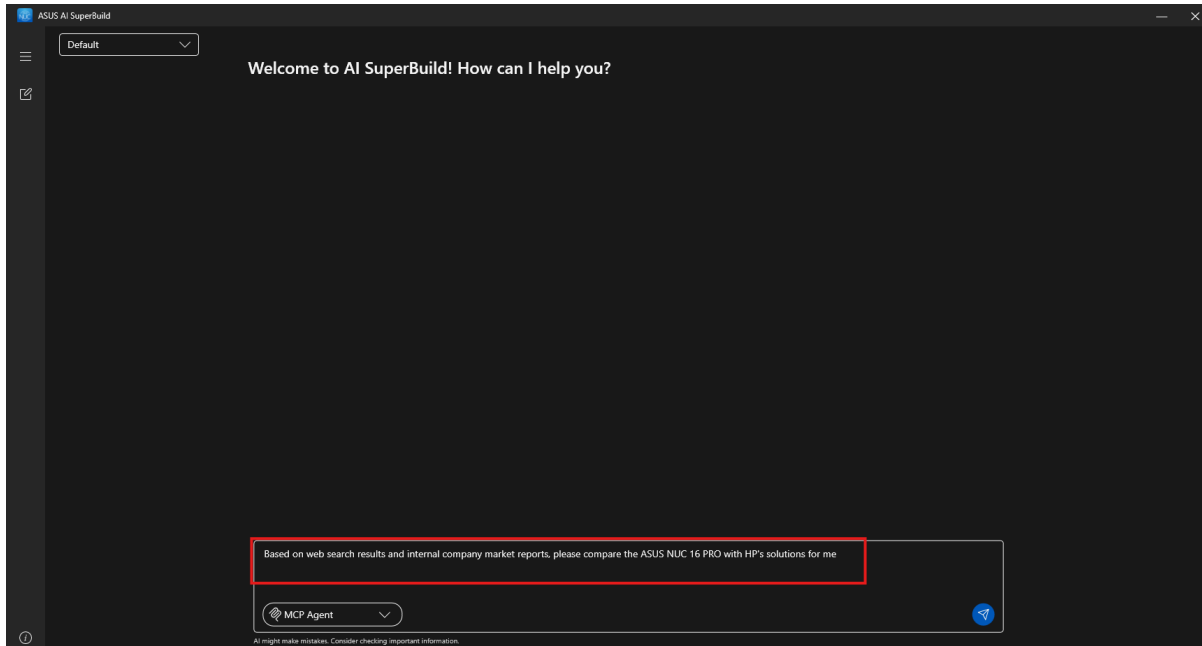
- b. Click the Back arrow at the top to return to the Home page.



c. Toggle the switch to Enable Brave Search



- d. Start the Demo: You can now ask the AI questions that require live web access.
- Demo Question Example: Type "Based on web search results and internal company market reports, please compare the ASUS NUC 16 PRO with HP's solutions for me." and press Enter. The AI will pull real-time data to generate the comparison.

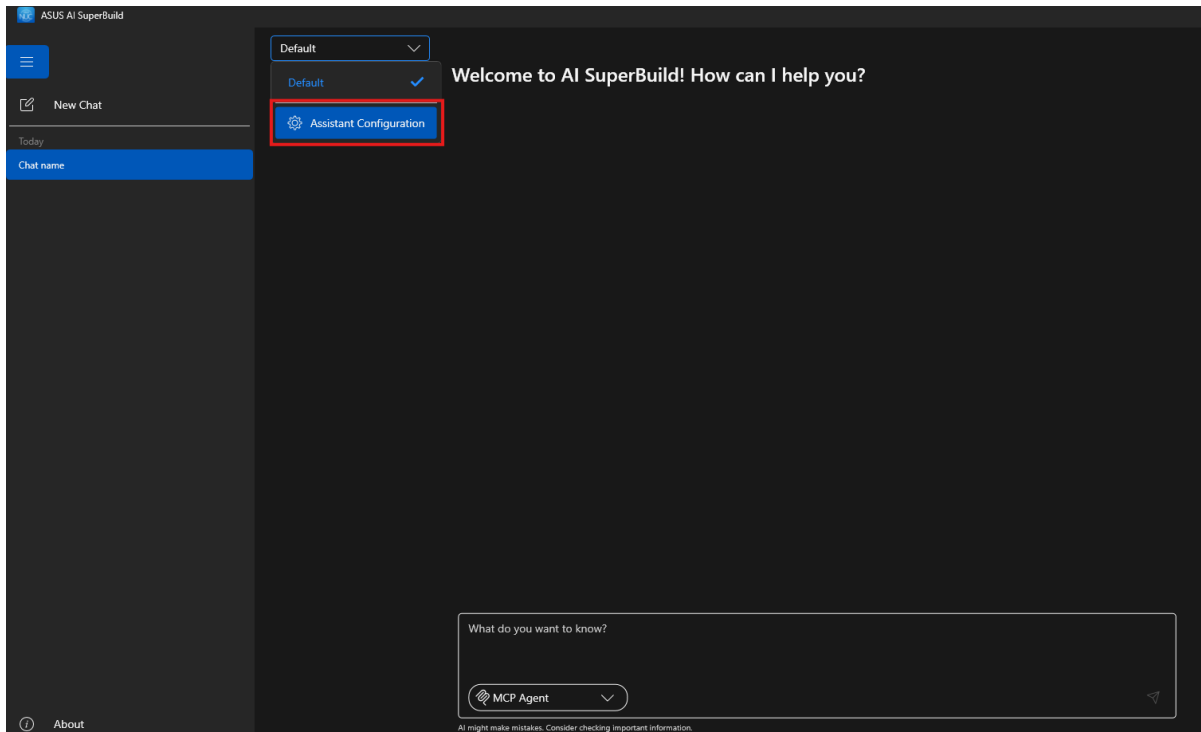


C. Scenario 3: Automated Presentation Generation (PDF to PPT)

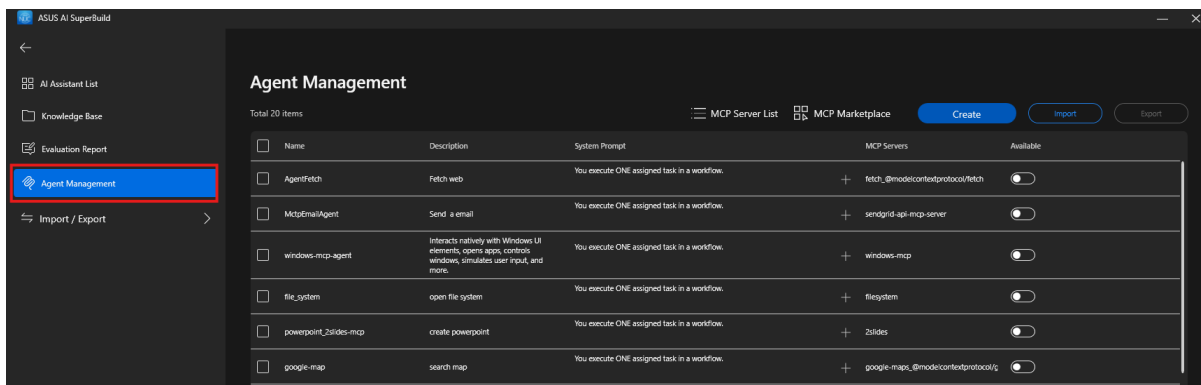
This advanced scenario demonstrates how the ASUS NUC AI Super Build can utilize multiple MCP tools simultaneously. The AI will access a local file drive, read a specific PDF document, and automatically generate a PowerPoint presentation based on its contents.

A. Access Server Configuration

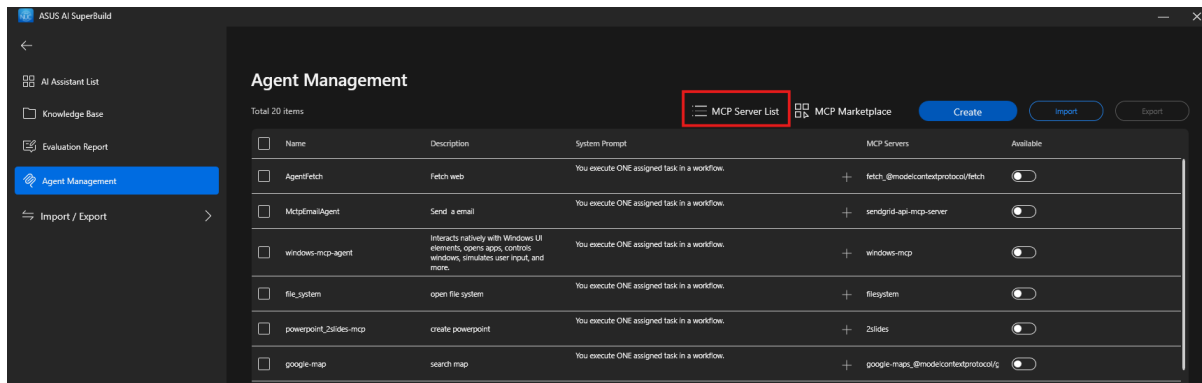
- a. On the main application interface, click Assistant Configuration.



- b. Select Agent Management.

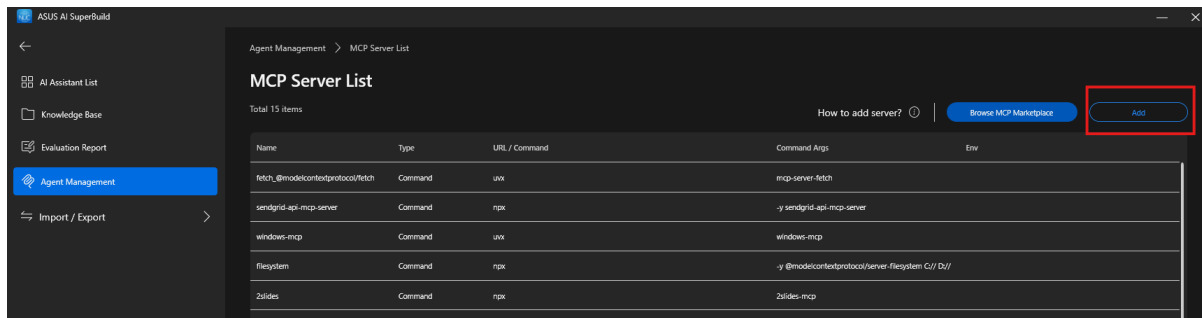


c. Click the MCP Server List tab.



B. Step 2: Add the Three MCP Servers

a. Click the Add button to configure a new server.



b. You will need to create three separate servers. For each one, click Add, set the Type dropdown to Command, input the exact details from the table below, and click Submit.

MCP Server Name	MCP Server Command	MCP Server Command Args (optional)	MCP Server Env (optional)
filesystem	npx	-y @modelcontextprotocol/server-filesystem C:// D://	(Leave Blank)
pdf-reader	npx	@sylphx/pdf-reader-mcp	(Leave Blank)
2slides	npx	2slides-mcp	(Leave Blank)

Filesystem:

Edit MCP Server

Type

URL Command

MCP Server Name

MCP Server Command

MCP Server Command Args (optional)

MCP Server Env (optional)

PDF-reader:

Edit MCP Server

Type

URL Command

MCP Server Name

MCP Server Command

MCP Server Command Args (optional)

MCP Server Env (optional)

2 Slides:

Edit MCP Server

Type

URL Command

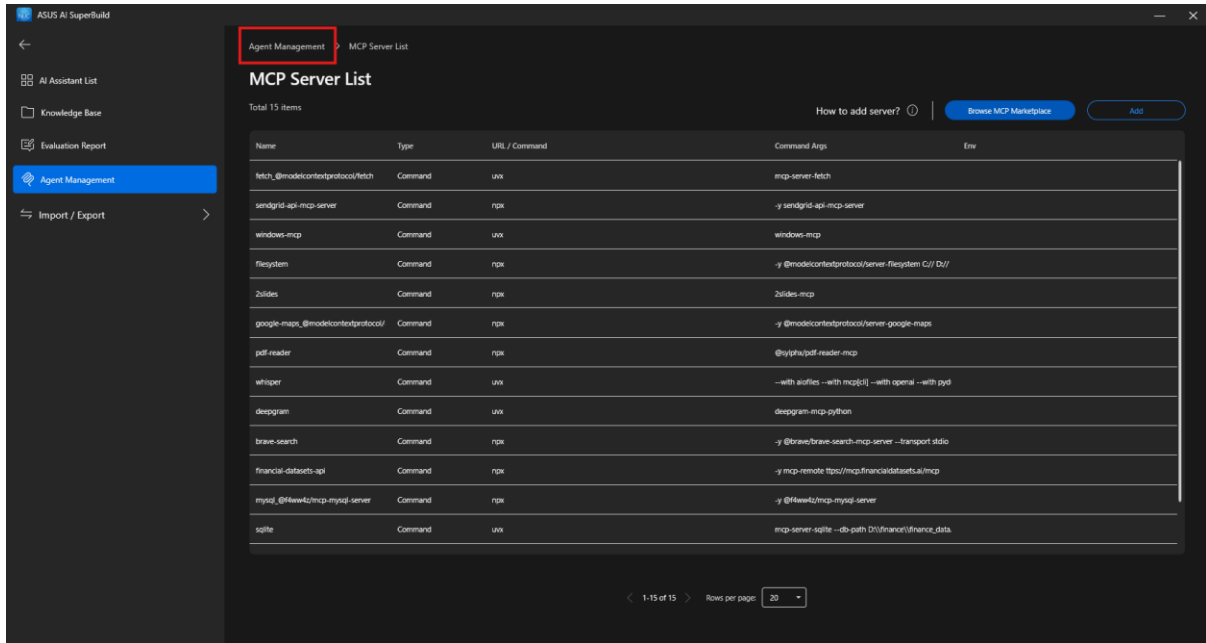
MCP Server Name

MCP Server Command

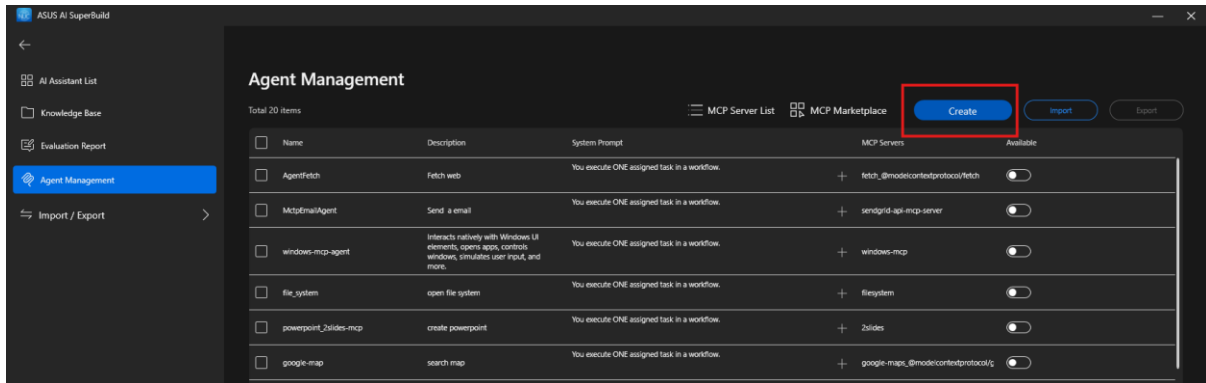
MCP Server Command Args (optional)

MCP Server Env (optional)

- C. Step 3: Create the Three MCP Agents
 - a. Return to the main Agent Management page.



- b. Click the Create button to create three corresponding agents.



c. For each one, fill in the details from the table below, and click Submit.

MCP Agent Name	Description	System Prompt	MCP Servers (Select Checkbox)
file_system	open file system	(no modification is needed)	filesystem
agent_pdf-reader	read pdf	(no modification is needed)	pdf-reader
powerpoint_2slides-mcp	create powerpoint	(no modification is needed)	2slides

File system:

Edit MCP Agent

MCP Agent Name

Description

System Prompt

You execute ONE assigned task in a workflow.

INPUTS:

MCP Servers

filesystem
×

Submit

Cancel

agent_pdf-reader:

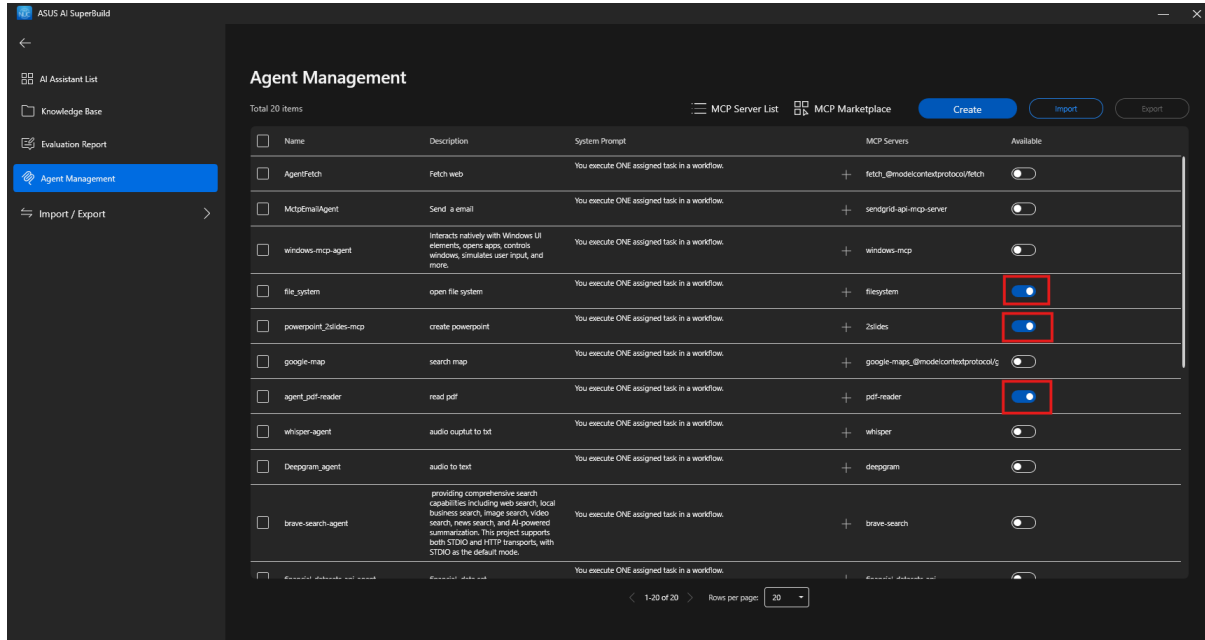
The screenshot shows a dark-themed 'Edit MCP Agent' form. At the top, the title 'Edit MCP Agent' is displayed in blue. Below it, there are four main sections: 'MCP Agent Name' with a text input containing 'agent_pdf-reader'; 'Description' with a text input containing 'read pdf'; 'System Prompt' with a text area containing 'You execute ONE assigned task in a workflow.' and 'INPUTS:'; and 'MCP Servers' which includes a blue pill-shaped tag 'pdf-reader' with a close icon and a dropdown menu below it containing '+ Select MCP server(s)'. At the bottom, there are two buttons: a blue 'Submit' button and a white 'Cancel' button with a blue border.

powerpoint_2slides-mcp:

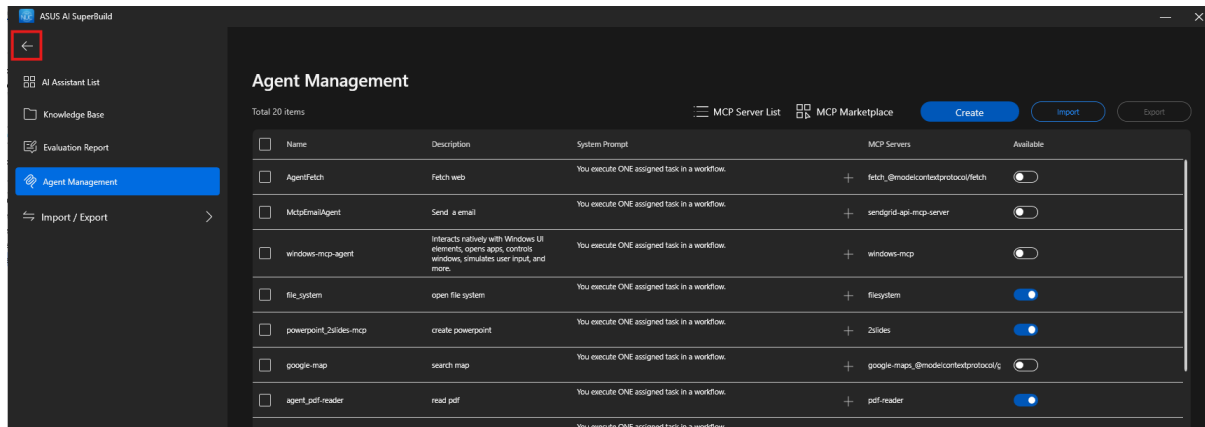
The screenshot shows a dark-themed 'Edit MCP Agent' form. At the top, the title 'Edit MCP Agent' is displayed in blue. Below it, there are four main sections: 'MCP Agent Name' with a text input containing 'powerpoint_2slides-mcp'; 'Description' with a text input containing 'create powerpoint'; 'System Prompt' with a text area containing 'You execute ONE assigned task in a workflow.' and 'INPUTS:'; and 'MCP Servers' which includes a blue pill-shaped tag '2slides' with a close icon and a dropdown menu below it containing '+ Select MCP server(s)'. At the bottom, there are two buttons: a blue 'Submit' button and a white 'Cancel' button with a blue border.

D. Step 4: Enable Agents and Execute the Demo

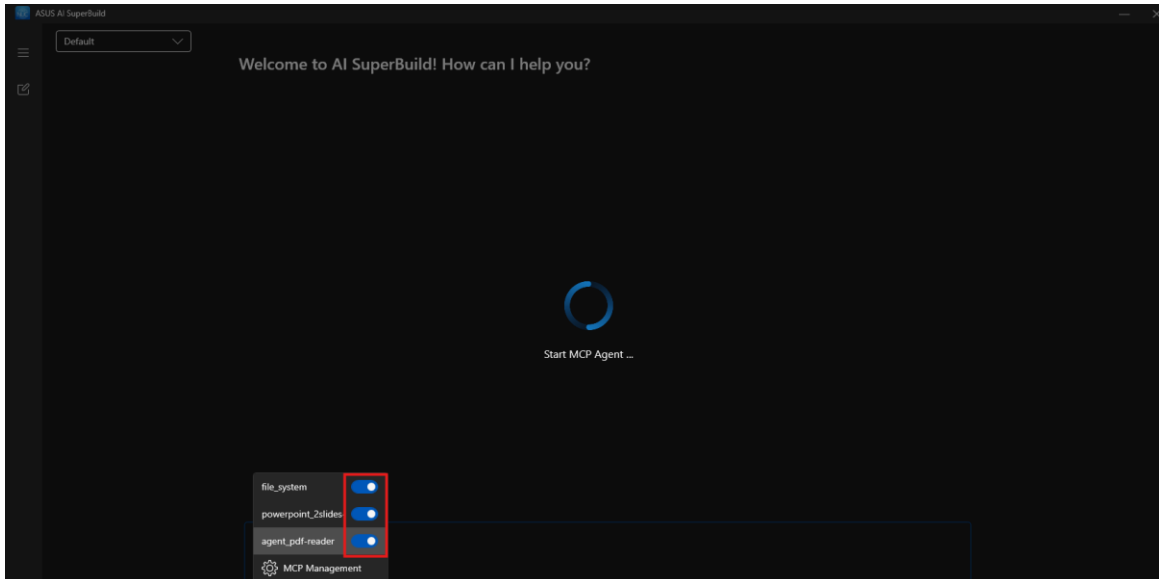
- a. Enable in Settings: While still on the Agent Management page, find your three newly created agents (file_system, agent_pdf-reader, powerpoint_2slides-mcp) and toggle their switches to Enable.



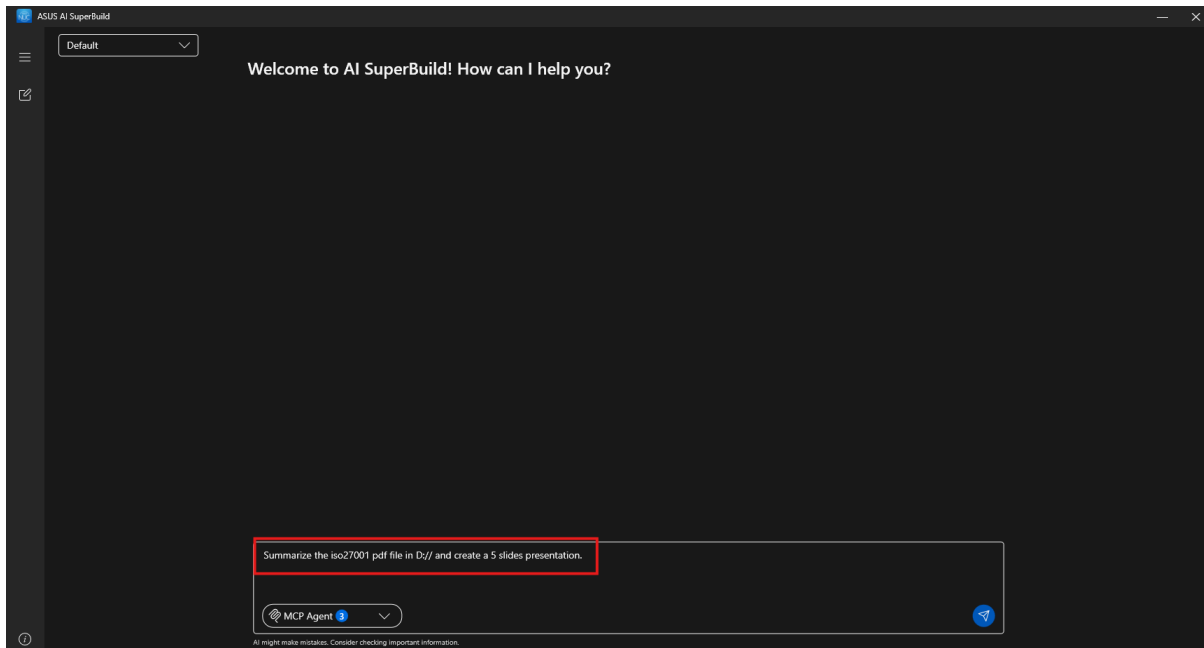
- b. Return to Home: Click the Back arrow at the top to return to the main chat interface.



- c. Activate for Chat: On the home page, ensure that the file_system, pdf-reader, and 2slides MCP agents are actively selected/turned on for your current session.



- d. Start the Demo: Copy and paste the following prompt into the chat and press Enter: "Summarize the iso27001 pdf file in D:// and create a 5 slides presentation."



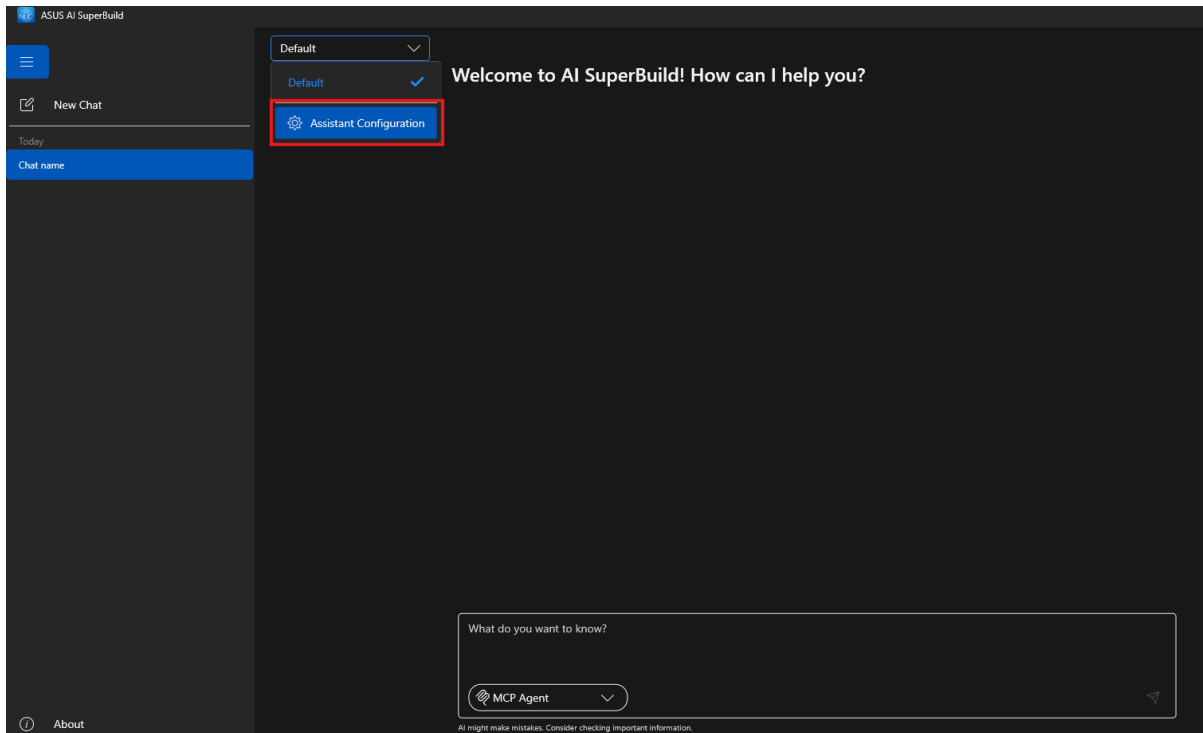
- e. Allow the AI a few moments to read the local drive, process the document, and output the finished slide deck for your audience.

D. Scenario 4: Financial Report Anomaly Detection (SQLite Database Analysis)

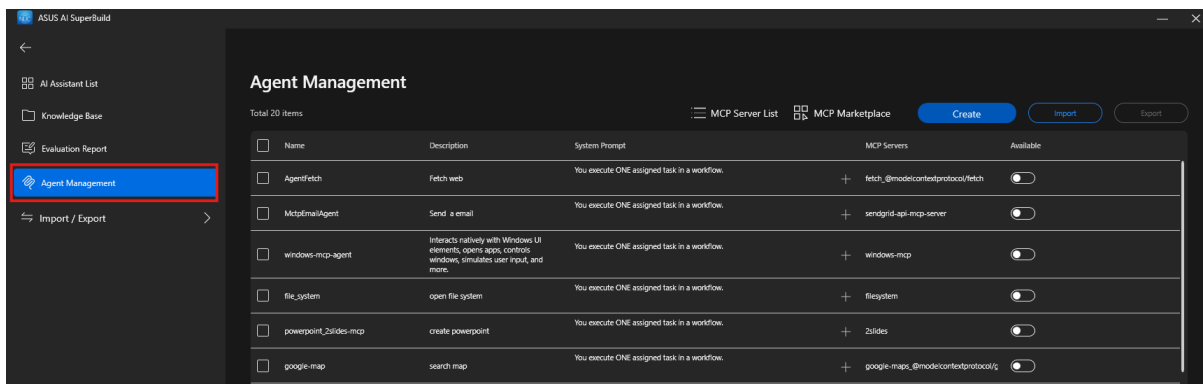
This scenario demonstrates how the AI assistant can connect directly to a local database file using the SQLite MCP to automatically cross-reference data across months and detect financial anomalies.

A. Step 1: Add the MCP Server

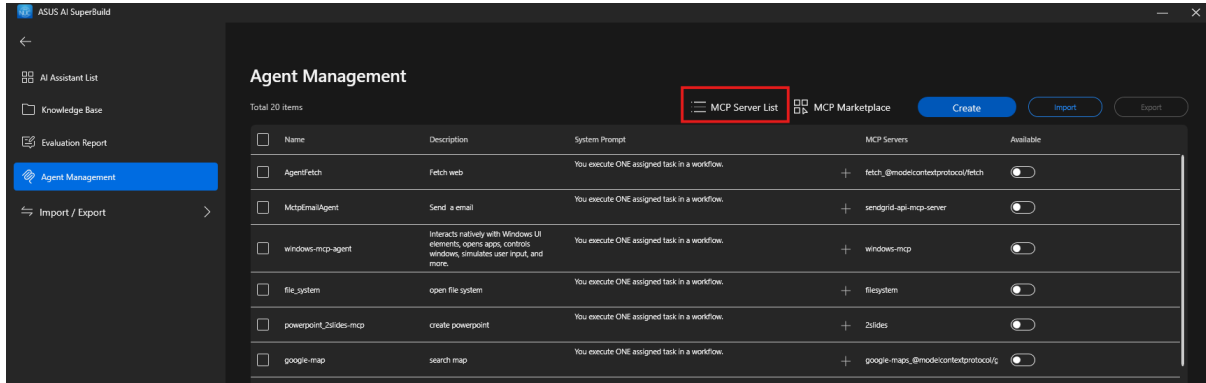
- a. On the main application interface, click Assistant Configuration.



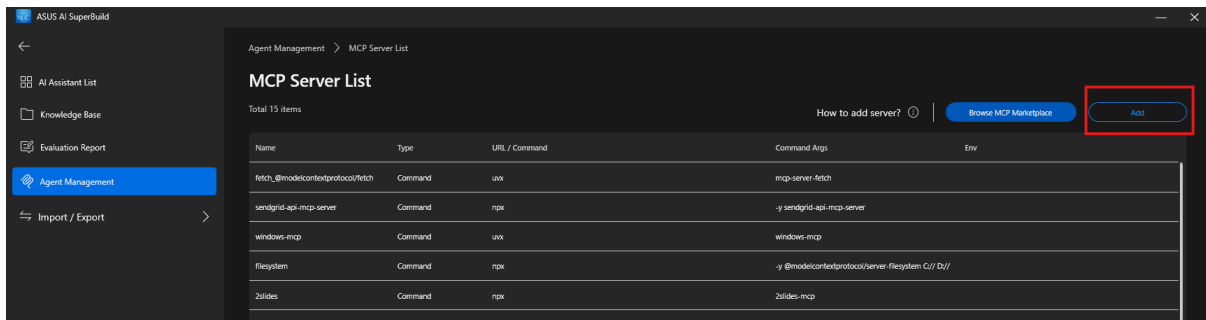
- b. Select Agent Management.



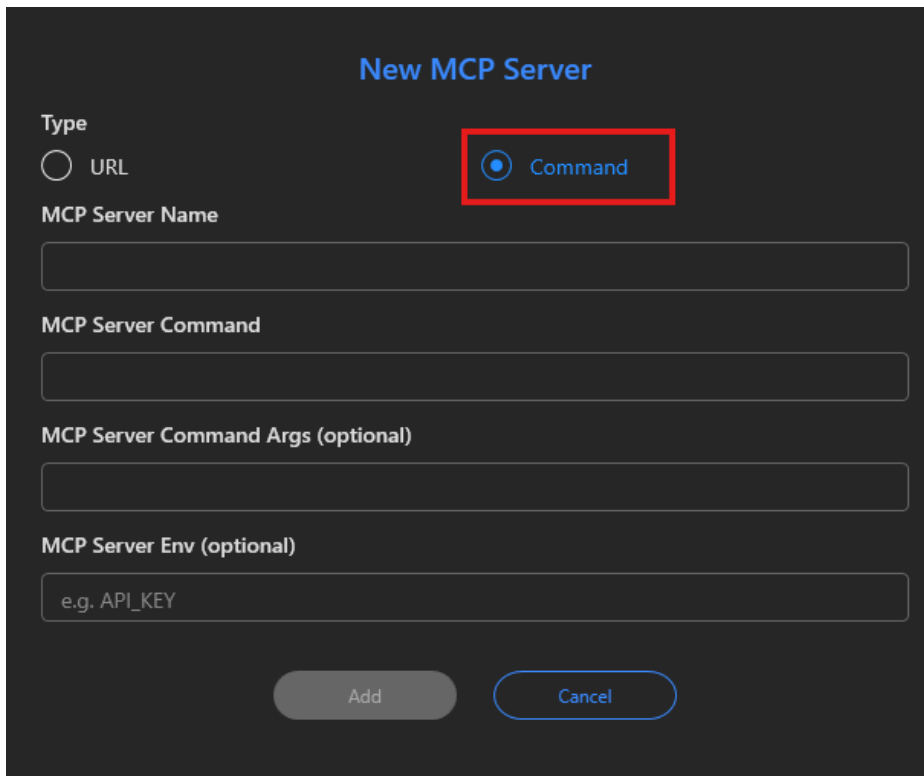
c. Click the MCP Server List tab.



d. Click the Add button to configure a new server.



e. In the configuration window, set the menu to Command.



f. Fill out the form using the exact text below, then click Submit:

- MCP Server Name: sqlite
- MCP Server Command: npx
- MCP Server Command Args (optional): mcp-server-sqlite --db-path D:\finance\finance_data.db

(⚠ Note: Please confirm the actual location of the .db file on your specific demo device and update this file path if necessary.)

- MCP Server Env (optional): (Leave Blank)

Edit MCP Server

Type

URL Command

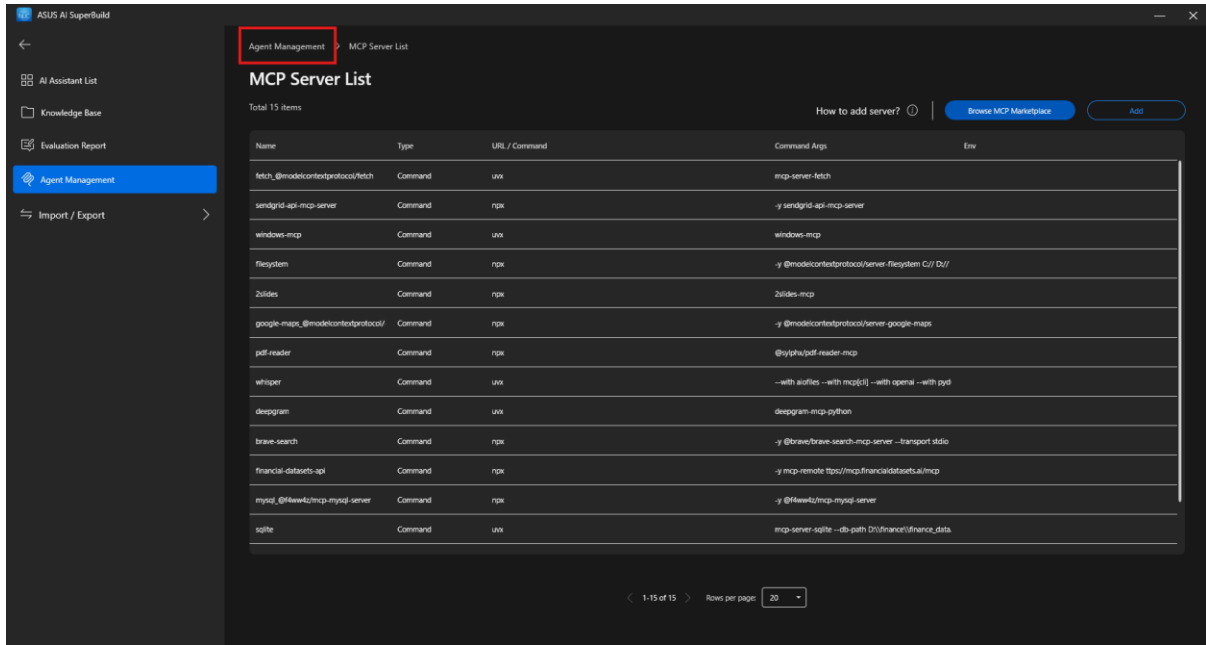
MCP Server Name

MCP Server Command

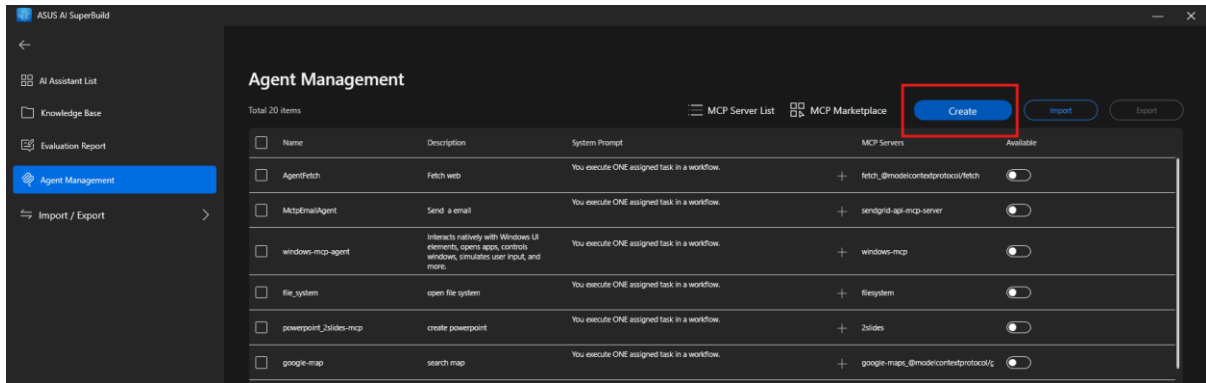
MCP Server Command Args (optional)

MCP Server Env (optional)

- B. Step 2: Create the MCP Agent
 - a. Return to the main Agent Management page.



- b. Click the Create button.



- c. Fill out the Agent details as follows, then click Submit:
- MCP Agent Name: sqlite
 - Description: database
 - System Prompt: (no modification is needed)
 - MCP Servers: Check the box for sqlite (the server you just created in Step 1).

Edit MCP Agent

MCP Agent Name

Description

System Prompt

You execute ONE assigned task in a workflow.

INPUTS:

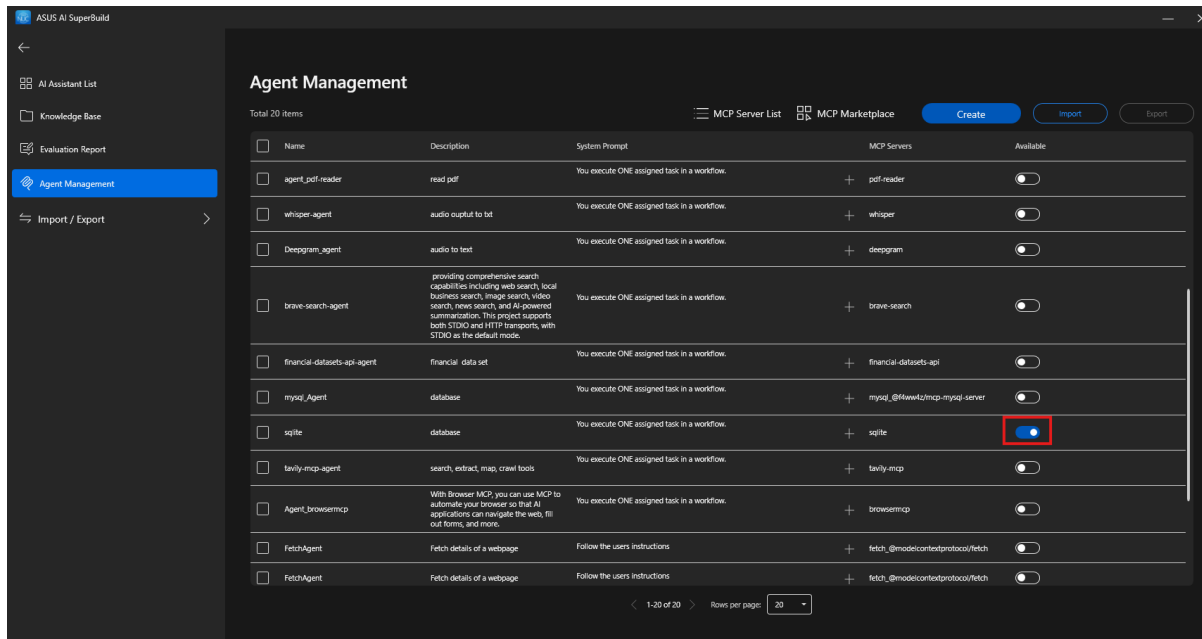
MCP Servers

sqlite X

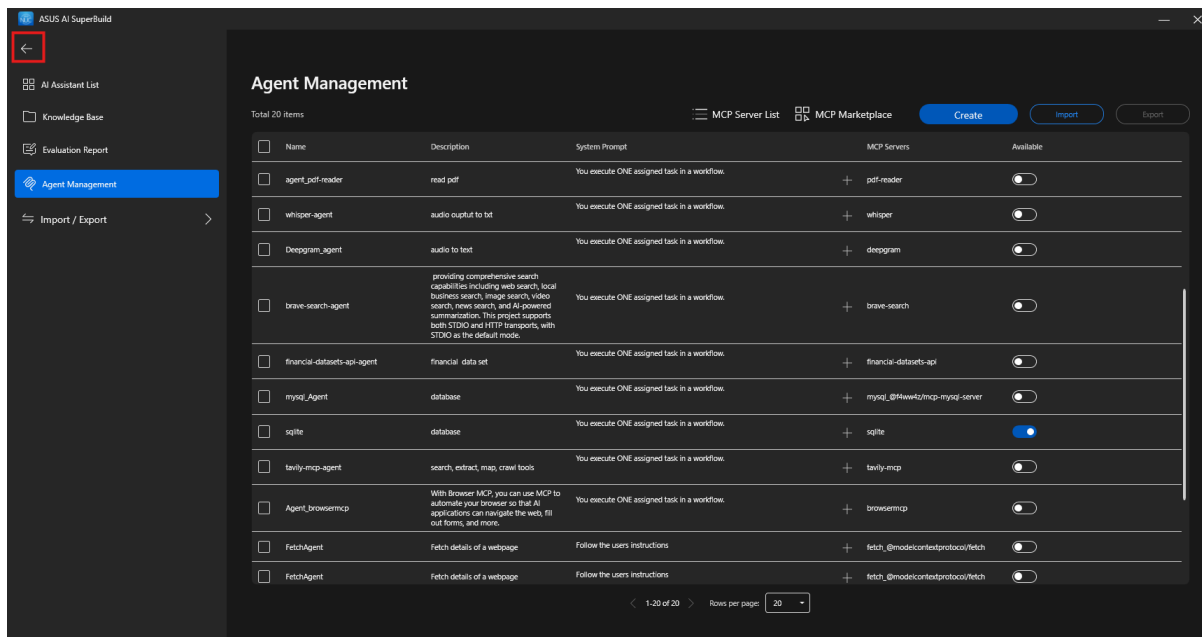
Submit Cancel

C. Step 3: Enable and Execute the Demo

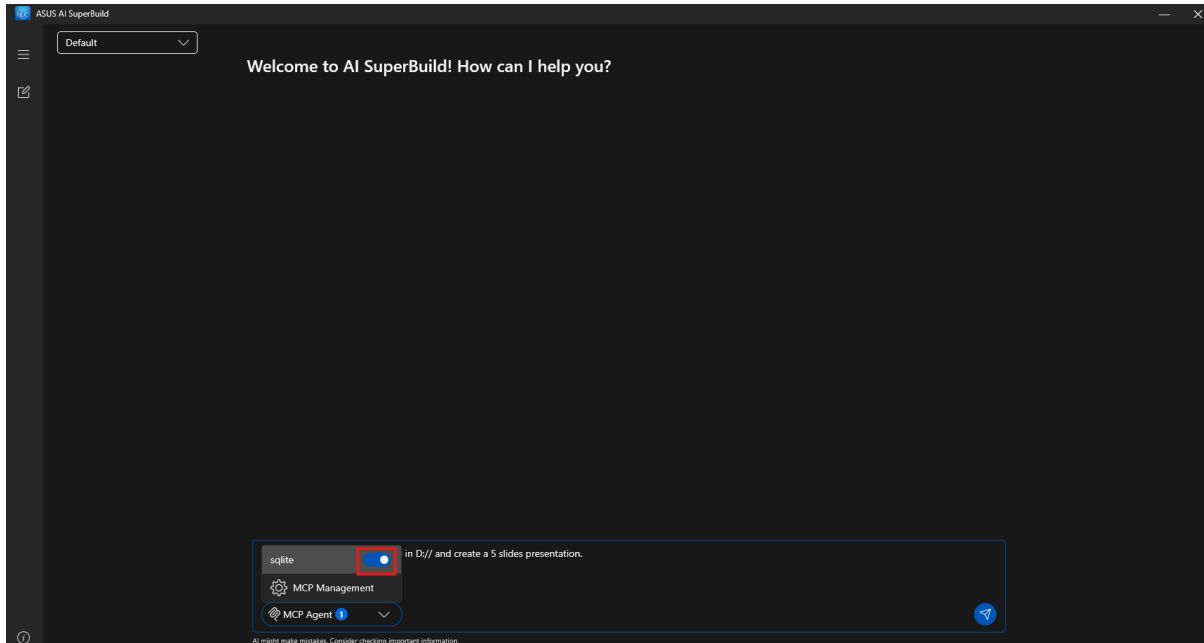
- a. Locate the sqlite agent you just created in the list and toggle the switch to Enable.



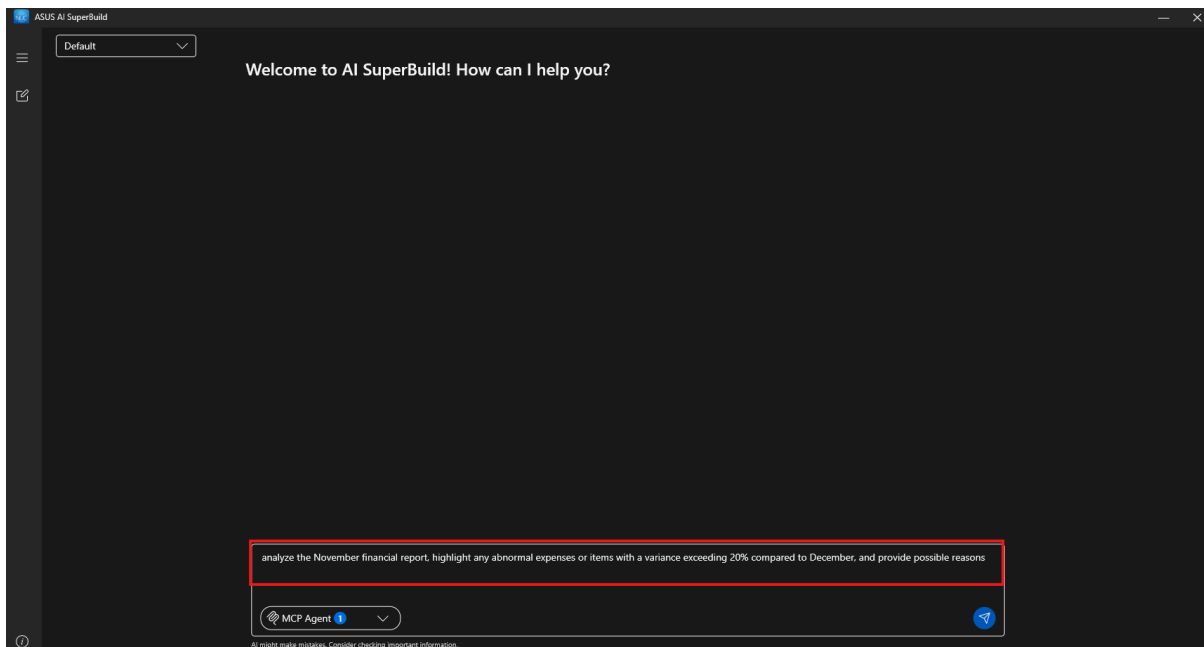
- b. Click the Back arrow at the top left to return to the Home page.



- c. Activate for Chat: On the home page, ensure that the the sqlite agent is actively selected for your current session.



- d. Start the Demo: Copy and paste the following prompt into the chat and press Enter: "analyze the November financial report, highlight any abnormal expenses or items with a variance exceeding 20% compared to December, and provide possible reasons"



- e. Allow the AI a few moments to read the local database and generate the analysis, demonstrating its ability to catch abnormal expenses and infer business context.