

## Build Your Own Custom Copilot Using Microsoft Copilot Studio!

# Agenda

Introduction to Microsoft Copilot Studio

Lab 1: Create your first copilot in Microsoft Copilot Studio

Lab 2: Authoring 101

Lab 3: Knowledge sources, AI knowledge, and custom instructions

Lab 4: Build and Invoke Power Automate cloud flows

Lab 5: Invoke AI Builder prompts

Lab 6: Make HTTP requests to connect to an API

Lab 7: Use generative AI orchestration to interact with your connectors

Lab 8: Using topic inputs

Q&A

# Introduction to Copilot Studio

# What is an agent?

An experience using generative AI to assist humans with complex cognitive tasks.

Chat using natural language (and code)

You determine, guide, and approve the output

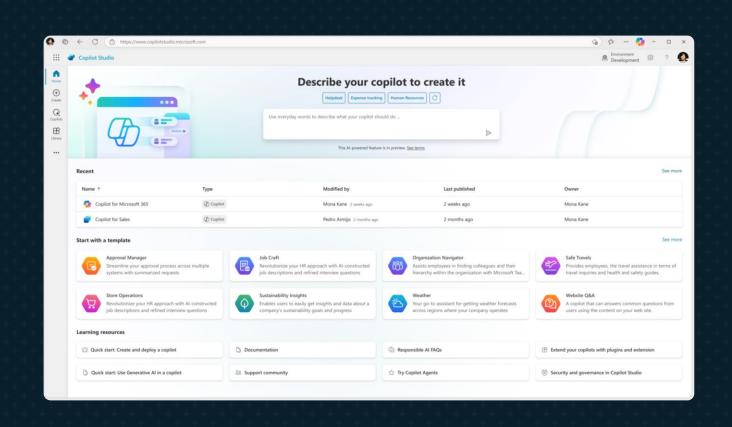
Its value increases with complexity



# Microsoft Copilot Studio

Your agent, your way

Copilot Studio is an end-to-end conversational AI product for building your own agent or **extending Microsoft Copilot** with generative Al, large language models and your data.



#### Create, manage, publish and extend Copilots

Live in minutes - all from one tool and E2E SaaS service

Integrate with Al Services
Integrate with Azure Al Studio,
Azure Cog Services, Bot
Framework and various other
Microsoft conversational services



## **Build & Publish**

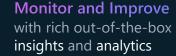


**Analyze & Improve** 



#### Chat over knowledge with Gen Al

Get enterprise specific answers over your files, websites, internal shares, Dataverse, third party systems and more



Publish to multiple channels,

and go live instantly on the SaaS

service or choose to extend

your custom capabilities

Copilot for Microsoft 365 with



#### Create specific topics

Supplement generative responses with specific, curated topics where you want tight control. Build them easily with the powerful graphical studio



#### **Build actions & Plugins**

Create actions, plugins, use 1000s of pre-built connectors or Power Automate to call your backends and APIs





# Copilot Studio message capacity add-on

\$200 per tenant/month
Tenant-wide capacity license sold in capacity packs.
1 capacity pack = 25,000 messages/month

#### **Offer Details**

- Gen Al: Al-enabled conversations
- Copilot in Microsoft Copilot Studio: Create and iterate on chatbot topics using expanded natural language capabilities
- Power Automate for Microsoft Copilot Studio: Cloud flows Create automated, instant and scheduled flows within context of Microsoft Copilot Studio bots
- Connectors: <u>Standard</u>, <u>Premium</u> and <u>Custom</u> connectors
- Dataverse for Microsoft Copilot Studio: Dataverse Default Capacity
  - 5 GB Database Capacity
  - 20 GB File Capacity
  - 2 GB Log Capacity

## **Get Started**



# Go to aka.ms/PPCCMCSWorkshop



**SCAN ME** 

### THEN:

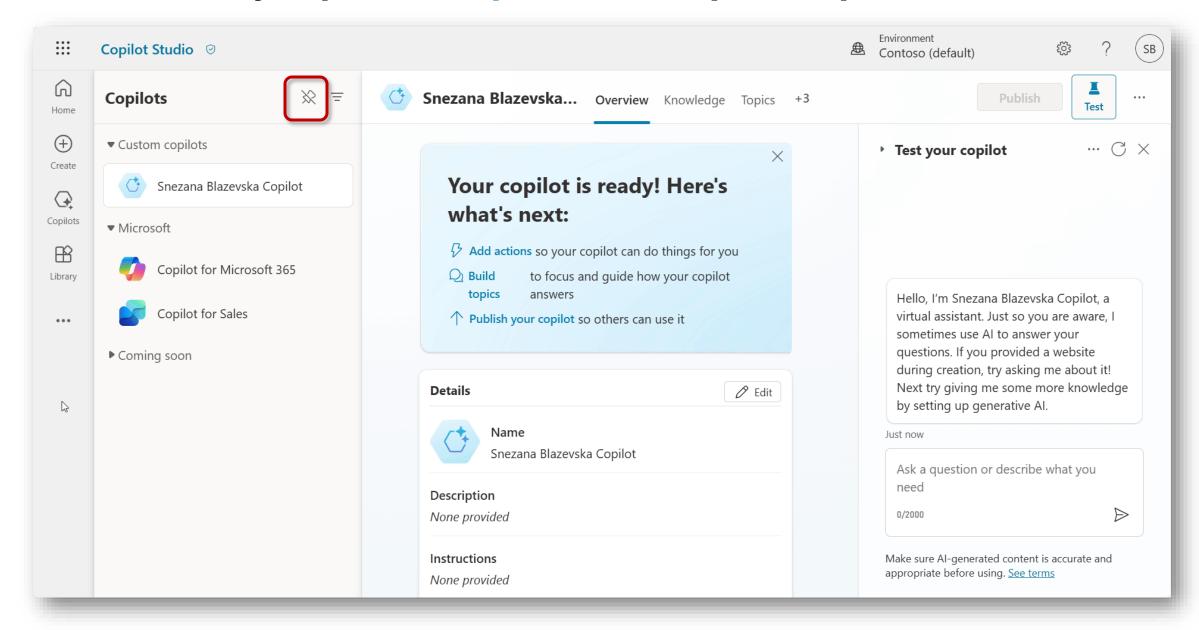
- **1. Download** and **unzip** the lab files locally.
- 2. Using the provided credentials, log in aka.ms/CopilotStudioStart

## IMPORTANT LAB INSTRUCTIONS

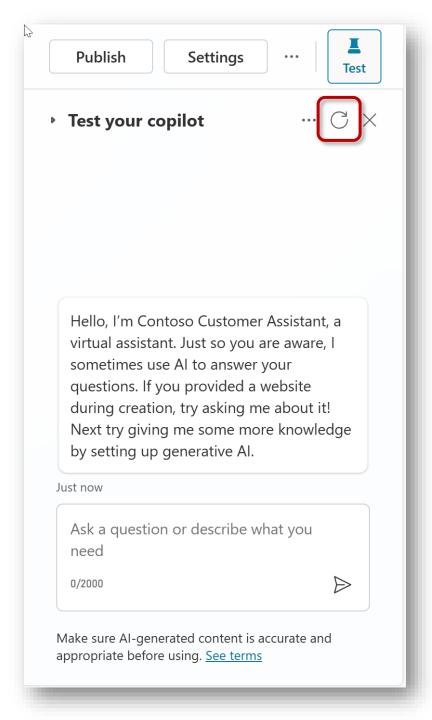
1. Follow the labs most lab content is dependent on previous labs.

2. If you want to deviate from the labs and test your own scenarios, this is fine, but please do complete the labs first, and create your own separate custom copilot for your own scenarios and exploration after you've completed the labs.

## Productivity tip #1: Unpin the Copilots pane



# Productivity tip #2: Refresh the conversation between tests



# Lab 01: Create your first copilot in Microsoft Copilot Studio

## Demo Lab 1

Test your copilot

.. C ×

Hello, I'm Contoso Copilot, a virtual assistant. Just so you are aware, I sometimes use AI to answer your questions. If you provided a website during creation, try asking me about it! Next try giving me some more knowledge by setting up generative AI.

Just now

Ask a question or describe what you need

0/2000



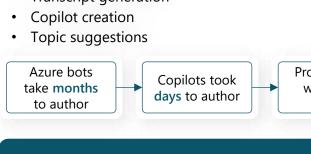
Make sure Al-generated content is accurate and appropriate before using. <u>See terms</u>

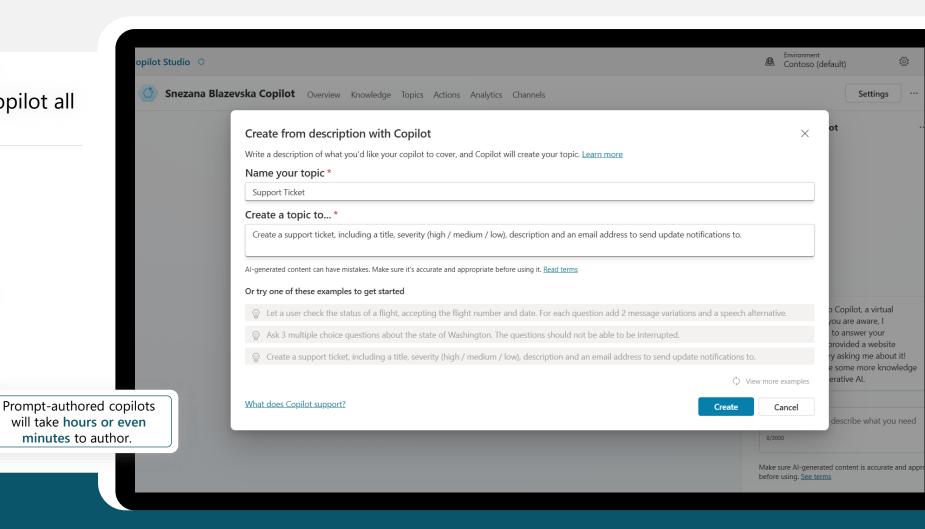
## **Copilot for Copilot Studio**

Al assistance in building topics, designing and modifying the copilot all through natural language

#### Assisted authoring for:

- Topic creation
- Topic iteration
- Response generation
- Adaptive Card generation
- Topic improvement suggestions
- Suggested trigger phrases, topic names, topic descriptions
- Transcript generation





# Lab time!

**Lab 01 [40] minutes** 

# Lab 02: Authoring 101

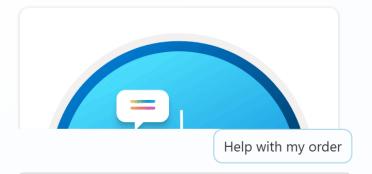
## Demo Lab 2





Hello, I'm Contoso Copilot, a virtual assistant. Just so you are aware, I sometimes use AI to answer your questions. If you provided a website during creation, try asking me about it!

Next try giving me some more knowledge by setting up generative AI.



Ask a question or describe what you need

0/2000



Make sure Al-generated content is accurate and appropriate before using. <u>See terms</u>

## Copilot Studio authoring canvas



#### Rich Responses

New features including:





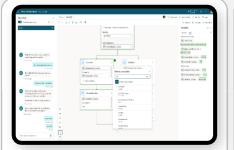
**Images** 

Adaptive Cards



Videos

**Quick Replies** 



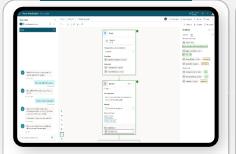
## Power Fx-based variables system

- Power Fx integrated formula editor for variable management
- List / Record type support
- Condition nodes with Power Fx formula support
- Iterate over a List with new Loop nodes (coming soon)



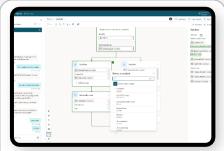
## Code side-by-side

- Move between a visual editor and code
- Share and re-use topic logic between developers and bots



## Event-driven & contextual triggering

- Trigger an event that can interrupt Dialogs if it detects a certain value has been received
- Add more nuanced responses to a message when the bot runs into an issue.
- Bots can stay within context when a user asks for additional help or more details



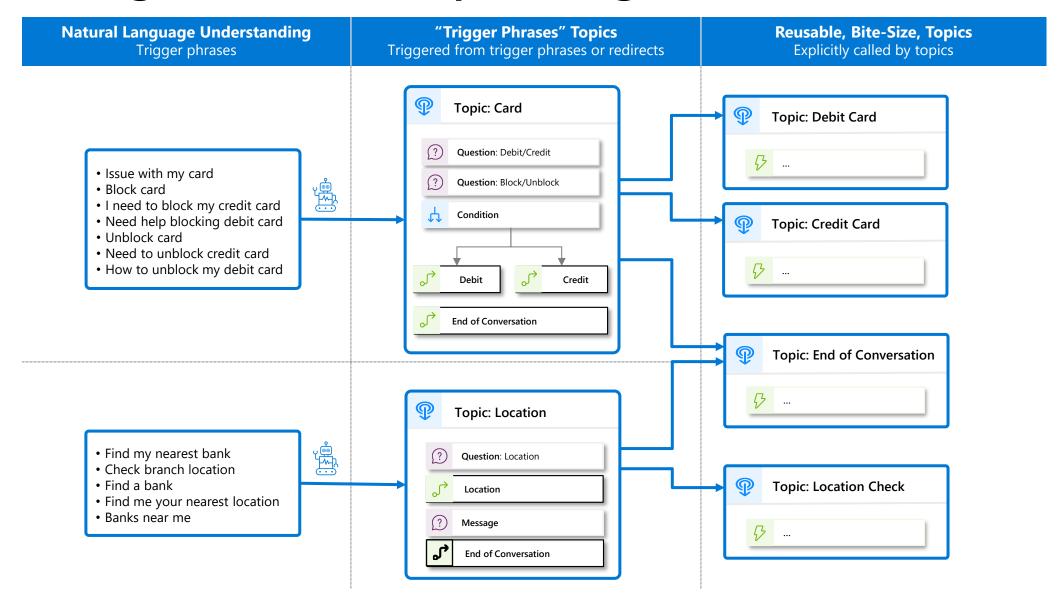
#### **Productivity**

- Cut, Copy, Paste
- Multi-node selection
- Enhanced topic navigation

# Lab time!

Lab 02 [60] minutes

## Disambiguation with topic design



## Natural language understanding

	Standard NLU model	Custom Azure CLU model	Generative orchestration
<b>⊘</b>	<ul> <li>✓ Default, out-of-the-box, model that comes pre-trained, with many predefined entity types.</li> <li>✓ Configuration is done by adding trigger phrases and custom entities (either closed lists with values and synonyms, or regular expressions).</li> </ul>	<ul> <li>✓ Supports additional languages, with native models.</li> <li>✓ Allows to further customize the intent triggering model for better intent recognition or to address specific industry requirements.</li> <li>✓ Allows for complex entity extraction (e.g., of the same type).</li> <li>✓ Entity extraction can also leverage Copilot Studio standard NLU.</li> </ul>	<ul> <li>✓ Uses a GPT large language model.</li> <li>✓ Can handle multiple intents and chain topics and/or plugins.</li> <li>✓ Automatically generate questions for missing inputs and answers complex entities and questions from the conversation context.</li> <li>✓ Configuration is done by simply describing topics, plugin actions, and inputs/outputs.</li> </ul>
<b>⊘</b>	<ul> <li>✓ Single intent recognition per query.</li> <li>✓ Cannot be extended.</li> <li>✓ Slot-filling multiple entities of the same type in the same query requires disambiguation for each (e.g., from and to cities)</li> </ul>	<ul> <li>✓ Single intent recognition per query.</li> <li>✓ Configuration is done in Azure and involves additional costs.</li> <li>✓ Has its own service limits that need to be evaluated.</li> <li>✓ Azure CLU intents and Copilot Studio topics must be carefully kept in sync.</li> </ul>	<ul> <li>✓ As it's a generative Al feature, the licensing burn rate of messages is higher than in regular topic triggering.</li> <li>✓ Preview capability.</li> </ul>

# Lab 03: Knowledge sources, AI knowledge, and custom instructions

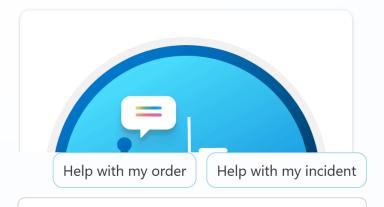
## Demo Lab 3

#### Test your copilot





Hello, I'm Contoso Copilot, a virtual assistant. Just so you are aware, I sometimes use AI to answer your questions. If you provided a website during creation, try asking me about it! Next try giving me some more knowledge by setting up generative Al.



Ask a question or describe what you need

0/2000



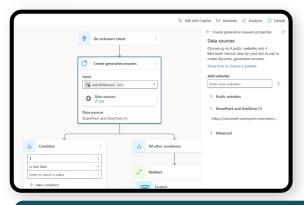
Make sure Al-generated content is accurate and appropriate before using. See terms

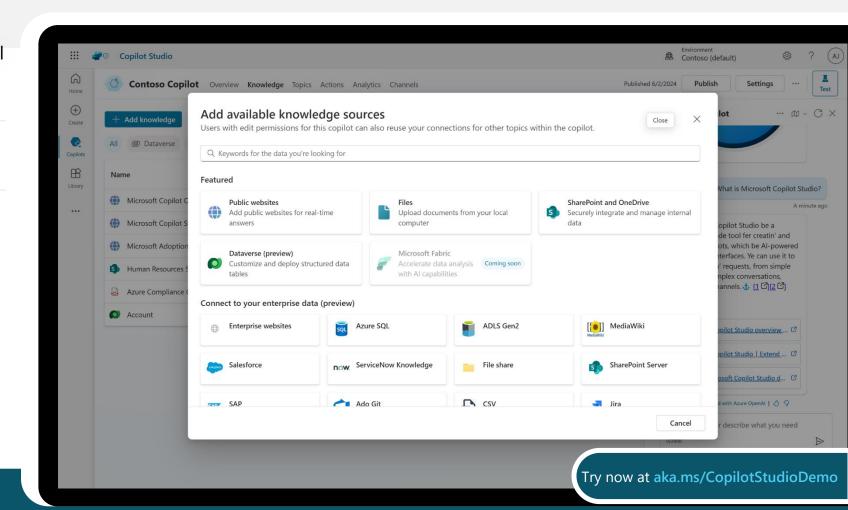
## **Knowledge sources**

Enable multi-turn chat over your own internal and external knowledge sources and sites with generative answers

Copilots can answer thousands of questions out of the box in seconds

You can even pull data from an API or other backend system and enable generative chat over it





Knowledge sources & Generative Al

**Query rewriting** 

Optimizing the user

question for search

Last 10 turns are leveraged

for contextualization

Querying each source

Top 3 results per source

Answer summarization from retrieved content

Citation generation

Answer personalization

with custom instructions

Content retrieval

**Summarization** 



#### **Public data**

- Websites must be indexed by Bing.
- Bing being a global service, no regional boundary is possible.

Website

Max 2 subpages depth (/en/help/), no direct pages.

### Bing Search

- One configuration ID, but it can be set using a formula. Azure costs covered by Microsoft
- Custom Up to 400 URLs, custom ranking options.
  - Max 2 subpages depth (/en/help/), supports direct pages.



#### SharePoint

(A) Internal only

- Requires the user to be authenticated with Entra ID to make delegated calls to the Graph API.
- Matching files (max 3 MB) are retrieved to get detailed snippets to summarize.
- Security trimming: returned results only include content the user has read access to.
- Users eligible to Microsoft 365 Semantic Index benefit from better results.



#### **Azure OpenAl** on your data

- Returns results from the linked Azure Al Search index (keyword, semantic, and vector search).
- Not a passthrough to Azure OpenAI generative answers summarization and validation apply.
- Connection isn't delegated: no security trimming, no authentication requirement for the user.
- Can index content from public URLs/web addresses (useful if regional boundary is required).



#### **Files**

- Files (max 512 MB) are stored in Dataverse file storage and indexed with Dataverse Search.
- Not possible to offer a URL to the source documents in the citations.
- Documents are uploaded at the copilot level. Gen Answers node will search across all uploaded documents or none, when disabled at a node level.



#### **Custom data**

- Requires a step to query the data source using cloud flows, connectors, or HTTP requests.
- Results are passed as inputs to the generative answers to summarize an answer for the query.
- Input data needs to be in table format, with 3 properties: Content (typically snippets of relevant content), ContentLocation (optional, typically a URL) and Title (optional).



#### **■ Validation at each step**

Moderation of harmful, malicious, uncompliant, or copyrighted answers

> Grounding validation and hallucination removal



### **Dataverse**

- Dataverse tables (max 15).
- Natural language queries are transformed in analytical queries (e.g., SQL) over structured data



#### Graph **Connectors**

- Connect to additional enterprise knowledge sources that are indexed in Microsoft Graph.
- Salesforce, ServiceNow, Confluence, Azure DevOps Wiki, etc.

## Generative AI security and compliance considerations



The large language models (LLM) Copilot Studio depends upon are the foundation models trained by OpenAI.



Generative answers uses GPT-3.5 Turbo for summarization, but will soon be updated to GPT-40 as of Sep 2024.



Models are hosted on internal Azure OpenAI services, honoring the Microsoft Services Trust boundary.



Models are accessed and used following Microsoft Responsible Al principles and policies.



Custom instructions can be added to generative answers to influence format and filters.



Data storage and processing for Copilot Studio generative AI capabilities may result in data movement across regional boundaries. When local data storage and processing aren't available in the region, environment settings are available to disable features requiring Azure OpenAI or Bing Search, preventing data movement.



Copilot Studio does not collect, nor provide any customer data, for use in the training of LLMs.



For operations purposes only, conversations are stored temporarily in a service-operated and secured store so that authorized Microsoft employees via Secure Access Workstations (SAWs) with Just-In-Time (JIT) access requests may respond to support requests. Access can be controlled through <a href="Customer Lockbox">Customer Lockbox</a>.



Generative AI capabilities log additional troubleshooting data, but only when the thumbs up or down reactions are used by the copilot author in the Copilot Studio test pane.



Given the existing control layers and to avoid further logging of customer data, <u>Azure OpenAl abuse monitoring</u> is disabled for Copilot Studio generative Al features.

# Lab time!

Lab 03 [30] minutes

# Lab 04: Build and call Power Automate cloud flows from your copilot

## Demo Lab 4

#### Test your copilot





Hello, I'm Contoso Copilot, a virtual assistant. Just so you are aware, I sometimes use AI to answer your questions. If you provided a website during creation, try asking me about it!

Next try giving me some more knowledge

by setting up generative Al.



Just now

Help with my order

Help with my incident

Ask a question or describe what you need

0/2000



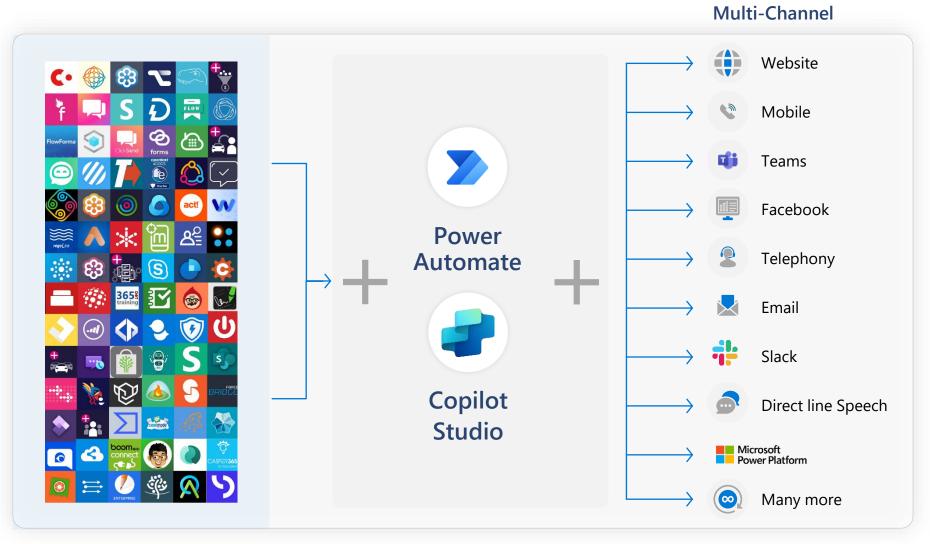
Make cure Al generated content is accurate and appropriate

## Integration patterns considerations

	Power Automate Cloud flows	HTTP requests & Connectors	Bot Framework Skills
<b>⊘</b>	<ul> <li>✓ No-code / low-code.</li> <li>✓ Clear separation of integration and conversational logic.</li> <li>✓ Can be monitored separately.</li> <li>✓ Existing cloud flows can be updated to integrate with Copilot Studio.</li> </ul>	<ul> <li>✓ No-code / low-code.</li> <li>✓ Faster runtime execution.</li> <li>✓ Can be monitored as part of Copilot Studio App Insights integration.</li> <li>✓ Can leverage variables, including environment variables and secrets.</li> <li>✓ Variables, conditions, parameters, etc. can use Power Fx formulas.</li> <li>✓ Parsing and error handling support.</li> </ul>	<ul> <li>✓ Synchronous execution.</li> <li>✓ Support for private endpoints.</li> <li>✓ Existing Bot Framework investments can be reused.</li> </ul>
<b>⊘</b>	<ul> <li>✓ Invocation and execution of the cloud flow can add latency.</li> <li>✓ If scale/performance are a concern, the higher-tier "Power Automate Process" plan can be evaluated.</li> <li>✓ "Power Automate Process" required for service principal ownership.</li> <li>✓ Need to return results within 100s.</li> </ul>	<ul> <li>✓ Mixing conversational logic and integration logic, but integration topics can be isolated, as they can be configured for inputs and outputs.</li> <li>✓ Need to return results within 100s.</li> </ul>	<ul> <li>✓ Pro-code (e.g., C#)</li> <li>✓ Runs in the Azure Al Bot Service.</li> <li>✓ Additional costs need to be covered by an Azure subscription.</li> <li>✓ ALM differs from Power Platform.</li> </ul>

### **Power Platform Connectors**

- Power Platform, including Power Automate and Copilot Studio, offers great integration capabilities, with more 1,200 native connectors or ways to create your own custom connectors to your APIs.
- For a good end-user experience, cloud flows, HTTP requests and connectors triggered from Copilot Studio must execute quickly so that the user doesn't have to wait too long for the copilot to answer.



1200+ prebuilt data connectors

# Lab time!

**Lab 04 [60] minutes** 

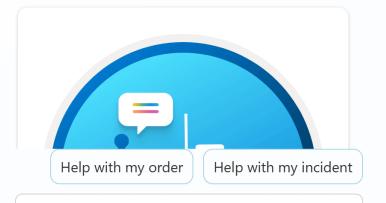
# Lab 05: Invoke AI Builder prompts

## Demo Lab 5

#### Test your copilot



Hello, I'm Contoso Copilot, a virtual assistant. Just so you are aware, I sometimes use AI to answer your questions. If you provided a website during creation, try asking me about it! Next try giving me some more knowledge by setting up generative Al.



Ask a question  $\phi$ r describe what you need

0/2000



Make sure Al-generated content is accurate and appropriate before using. See terms

# Lab time!

Lab 05 [25] minutes

## Lab 06: Make HTTP requests to connect to an API

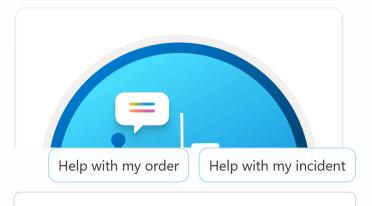
#### Demo Lab 6

#### Test your copilot





Hello, I'm Contoso Copilot, a virtual assistant. Just so you are aware, I sometimes use AI to answer your questions. If you provided a website during creation, try asking me about it! Next try giving me some more knowledge by setting up generative Al.



Ask a question or describe what you need

0/2000



Make sure Al-generated content is accurate and appropriate before using. See terms

## Lab time!

Lab 06 [20] minutes

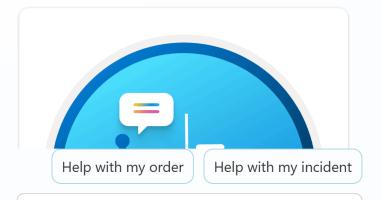
# Lab 07: Use generative AI orchestration to interact with your connectors

#### Demo Lab 7

#### Test your copilot



Hello, I'm Contoso Copilot, a virtual assistant. Just so you are aware, I sometimes use AI to answer your questions. If you provided a website during creation, try asking me about it! Next try giving me some more knowledge by setting up generative AI.



Ask a question or describe what you need

0/2000



Make sure Al-generated content is accurate and appropriate before using. <u>See terms</u>

#### **Generative AI orchestration**

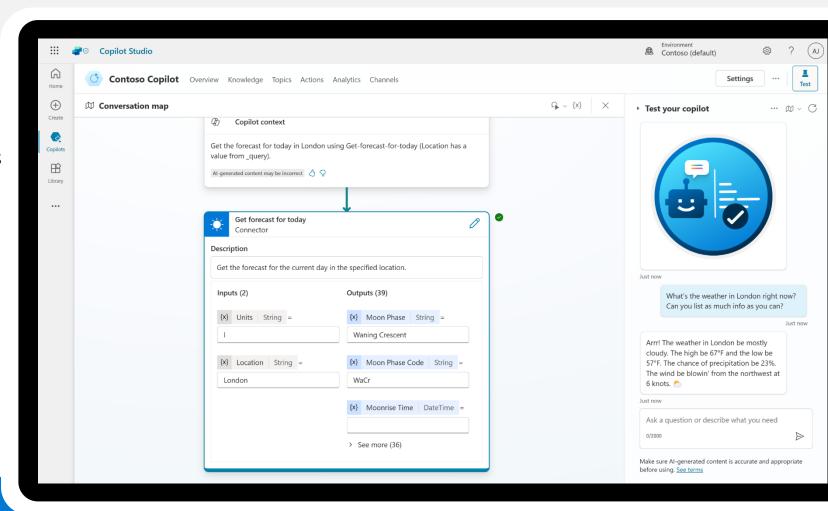
Test and trace feature to see it in action

Dynamically chains actions and topics together to produce actionable responses

Handle multi-intent queries that were not anticipated or built by the user

Automatically slot-fill further details to get the information needed for the task

Powered by the Azure OpenAl Service with LangChain concepts



## Lab time!

**Lab 07 [20] minutes** 

## Lab 08: Using topic inputs

## Lab time!

Lab 08 [15] minutes



## Q&A

### Resources

#### Resources



Copilot Studio trial <u>aka.ms/TryCopilotStudio</u>

Copilot Studio website <u>aka.ms/CopilotStudio</u>

Official Blog <u>aka.ms/CopilotStudioBlog</u>

Copilot Studio Demo <u>aka.ms/CopilotStudioDemo</u>

Copilot Studio Kit <u>aka.ms/CopilotStudioKit</u>

Product documentation <u>aka.ms/CopilotStudioDocs</u>

Product guidance <u>aka.ms/CopilotStudioGuidance</u>

Implementation guide <u>aka.ms/CopilotStudioImplementationGuide</u>

Community <u>aka.ms/CopilotStudioCommunity</u>